

**REGINALDO DE OLIVEIRA**

**OTIMIZAÇÃO DOS PESOS DAS OBSERVAÇÕES GEODÉSICAS  
CONSIDERANDO UMA MATRIZ CRITÉRIO DE COVARIÂNCIAS E O  
PROBLEMA DE VALOR PRÓPRIO INVERSO: APLICAÇÃO DOS  
ALGORITMOS DE OTIMIZAÇÃO MATEMÁTICA**

**Seminário II apresentado ao Curso de Pós-Graduação em Ciências Geodésicas, Departamento de Geomática, Setor de Ciências da Terra, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Ciências Geodésicas.**

**Orientador: Prof. Dr. Quintino Dalmolin**

**CURITIBA  
2005**

Oliveira, Reginaldo de

Otimização dos pesos das observações geodésicas considerando uma matriz critério de covariâncias e o problema de valor próprio inverso: aplicação dos algoritmos de otimização matemática / Reginaldo de Oliveira. — Curitiba, 2005.  
x, 114 f.: il.; tab.

Orientador: Quintino Dalmolin

Seminário apresentado ao Curso de Pós-Graduação em Ciências Geodésicas – Universidade Federal do Paraná. Setor de Ciências da Terra.

1. Ajustamento (Geodesia). 2. Redes (Geodesia). 3. Otimização matemática.
2. I. Dalmolin, Quintino.II. Título.

CDD 20 526.51

## AGRADECIMENTOS

A **Deus** pela força de cada dia. Agradeço a Ele pelas tristezas e alegrias;

A minha esposa Ângela Maria de Oliveira, sem ela a minha existência não estaria completa;

Ao meu pai Nélio Schoröder de Oliveira grande professor de minha vida;

A minha mãe Marlene de Oliveira pela vida;

A minhas irmãs Cleide e Neide pela paciência;

Ao meu orientador Prof. Dr. Quintino Dalmolin pelo incentivo que me deu para cursar o mestrado e posteriormente o doutorado dedicando-se e incentivando a minha formação. Agradeço de coração.

Ao colega Sandro Reginato Soares de Lima e Eddie Andretto Junior pelo auxílio nas horas que mais precisei.

A secretária do curso de pós-graduação em Ciências Geodésicas Verali Mônica Keleuser Reguilin, pela presteza de seus serviços.

Aos Professores Silvio Jacks dos Anjos Garnés e Luis Carlos Matioli pelas valorosas sugestões que enriqueceram este trabalho.

Aos docentes, funcionários e discentes do Colégio Estadual Dr. Munhoz da Rocha (distrito de Guaragi em Ponta Grossa) pelo companheirismo e dedicação ao trabalho.

Aos colegas do curso de Pós-Graduação em Ciências Geodésicas.

## CIDADEZINHA QUALQUER

Casas entre bananeiras  
mulheres entre laranjeiras  
pomar amor cantar.

Um homem vai devagar.  
Um cachorro vai devagar.  
Um burro vai devagar.

Devagar...as janelas olham.

Eta vida besta, meu Deus.

( Carlos Drummond de Andrade)

Quando entendi que podia morrer,  
pensei:  
não tem cabimento desperdiçar o resto  
da vida. Virei Albert Einstein, o  
defensor da relatividade: quando  
alguma coisa me desagrada, procuro  
avaliar que importância ela tem no  
universo.  
Descobri que é possível  
ser feliz até quando estou triste.

(Drauzio Varella)

## SUMÁRIO

<b>LISTA DE FIGURAS</b> .....	vi
<b>LISTA DE TABELAS E QUADROS</b> .....	vii
<b>LISTA DE SIGLAS</b> .....	vii
<b>LISTA DE SÍMBOLOS</b> .....	viii
<b>RESUMO</b> .....	ix
<b>ABSTRACT</b> .....	x
<b>1 INTRODUÇÃO</b> .....	1
1.1 CONSIDERAÇÕES SOBRE PROJETOS DE REDES GEODÉSICAS .....	3
1.2 HISTÓRICO .....	7
1.3 COOPERAÇÃO E SUPORTE PARA SUA REALIZAÇÃO .....	9
1.4 OBJETIVOS .....	10
1.4.1 Objetivo Geral .....	10
1.4.2 Objetivo Específico .....	10
1.5 CONTRIBUIÇÃO E ESTRUTURAÇÃO DA PESQUISA .....	11
<b>2 CONCEITOS FUNDAMENTAIS</b> .....	13
2.1 MATRIZ ORTOGONAL .....	13
2.2 INVERSA ORDINÁRIA .....	14
2.3 INVERSA GENERALIZADA E INVERSA GENERALIZADA DE MOORE- PENROSE .....	14
2.4 SOLUÇÃO DE UM SISTEMA DE EQUAÇÕES LINEARES PELOS MÍNIMOS QUADRADOS .....	16
2.5 PRODUTO DE KRONECKER .....	17
2.6 PRODUTO DE KHATRI-RAO .....	18
2.7 OPERADORES $\text{vec}$ e $\text{vecd}$ .....	19
2.8 VALORES E VETORES PRÓPRIOS: ESPECIAL E GENERALIZADO .....	20
<b>3 MÉTODOS ANALÍTICOS BASEADOS EM MATRIZ CRITÉRIO DE COVARIÂNCIAS PARA A OTIMIZAÇÃO DOS PESOS DAS OBSERVAÇÕES GEODÉSICAS</b> .....	23
3.1 PROBLEMA SEMI-DEFINIDO POSITIVO.....	24
3.1.1 Solução com o Produto de Kronecker .....	24
3.1.2 Solução para a matriz completa dos pesos com auxílio direto da Pseudo-Inversa....	25
3.2 PROBLEMA DIAGONAL .....	25
3.2.1 Aproximação Direta da Matriz Critério de Covariância .....	25
3.2.2 Aproximação Iterativa da Matriz Critério de Covariância .....	26
3.2.3 Aproximação Direta da Inversa da Matriz Critério de Covariância .....	27
3.3 EXEMPLIFICAÇÃO PARA OS MÉTODOS BASEADOS EM MATRIZ CRITÉRIO .....	28
<b>4 A OTIMIZAÇÃO DOS PESOS DAS OBSERVAÇÕES GEODÉSICAS BASEADA EM VALORES PRÓPRIOS</b> .....	34
4.1 CRITÉRIOS DE OTIMALIDADE PARA REDES GEODÉSICAS .....	36
4.2 TESTE PARA A IGUALDADE DE VALORES PRÓPRIOS .....	37
<b>5 O PROBLEMA DE VALOR PRÓPRIO INVERSO</b> .....	40
5.1 O PROBLEMA DE VALOR PRÓPRIO INVERSO ADITIVO .....	41
5.2 O PROBLEMA DE VALOR PRÓPRIO INVERSO MULTIPLICATIVO .....	41
5.3 O PROBLEMA DE MÍNIMOS QUADRADOS DE VALOR PRÓPRIO INVERSO .....	43

5.4 PROBLEMA DE VALOR PRÓPRIO INVERSO APLICADO NA OTIMIZAÇÃO DE PESOS EM LEVANTAMENTOS GEODÉSICOS.....	43
5.5 FORMULAÇÃO NUMÉRICA .....	45
5.5.1 Formulação I .....	45
5.5.2 Formulação II .....	45
<b>6 O PROBLEMA NÃO LINEAR</b> .....	47
6.1 ALGORITMOS .....	50
6.1.1 O Método de Newton .....	51
6.1.1.1 O Método de Newton para solução da formulação I .....	53
6.1.1.2 O método de Newton para solução da formulação II .....	55
6.2 OTIMIZAÇÃO:MINIMIZAÇÃO IRRESTRITA.....	58
6.2.1 Estratégias: Busca Linear e Região de Confiança .....	63
6.3 ALGORITMO DE DESCIDA COM BUSCA LINEAR .....	66
6.3.1 O Método de Newton Modificado com Busca Linear .....	69
6.4 O MÉTODO DA REGIÃO DE CONFIANÇA .....	72
6.5 MÉTODO QUASE NEWTON .....	78
6.6 MÉTODO LP PARA UM PROBLEMA DE VALOR PRÓPRIO INVERSO .....	80
<b>7 EXPERIMENTOS</b> .....	88
7.1 EXPERIMENTO 1 .....	89
7.2 EXPERIMENTO 2 .....	93
7.3 EXPERIMENTO 3 .....	98
<b>8 CONCLUSÃO</b> .....	103
<b>9 REFERÊNCIAS</b> .....	106
<b>APÊNDICE A – MATRIZ DEFINIDA E SEMI DEFINIDA POSITIVA</b> .....	110
<b>APÊNDICE B – NORMAS</b> .....	113

## LISTA DE FIGURAS

FIGURA 1.1	: ESTABELECIMENTO DE UMA REDE GEODÉSICA.....	4
FIGURA 3.1	: REDE DE NIVELAMENTO .....	29
FIGURA 4.1	: REPRESENTAÇÃO GRÁFICA DO ESPECTRO DE UMA MATRIZ .....	35
FIGURA 4.2	: REDE HOMOGÊNEA E ISOTRÓPICA .....	37
FIGURA 6.1	: MINIMIZADOR GLOBAL E LOCAL .....	49
FIGURA 6.2	: MÉTODO DE NEWTON: CONVERGÊNCIA .....	52
FIGURA 6.3	: POLINÔMIO CARACTERÍSTICO DE $A(c)$ .....	58
FIGURA 6.4	: COMPRIMENTO DO PASSO .....	64
FIGURA 6.5	: REGIÃO DE CONFIANÇA .....	65
FIGURA 6.6	: CONDIÇÃO DO ÂNGULO .....	67
FIGURA 6.7	: REGIÃO DE CONFIANÇA APROXIMADO .....	75
FIGURA 6.8	: COMPARATIVO: REGIÃO DE CONFIANÇA E O PASSO DE NEWTON.....	76
FIGURA 6.9	: GEOMETRIA DO MÉTODO LP .....	84
FIGURA 7.1	: REPRESENTAÇÃO GEOMÉTRICA DO PLANEJAMENTO DE UM PONTO EM COORDENADAS CARTESIANAS .....	90
FIGURA 7.2	: PROJETO DE REDES GEODÉSICAS: GEOMETRIA PARA OBTENÇÃO DE DOIS PONTOS EM COORDENADAS HORIZONTAIS .....	94
FIGURA 7.3	: REDE GEODÉSICA BIDIMENSIONAL .....	98

## LISTA DE TABELAS E QUADROS

TABELA 7.1	: COORDENADAS HORIZONTAIS I .....	91
TABELA 7.2	: RESULTADO DA OTIMIZAÇÃO: BUSCA LINEAR, REGIÃO DE CONFIANÇA, BFGS E LP .....	93
TABELA 7.3	: COORDENADAS HORIZONTAIS II .....	95
TABELA 7.4	: COORDENADAS HORIZONTAIS DE UMA REDE GEODÉSICA BIDIMENSIONAL .....	99
QUADRO 7.1	: RESULTADO DO PROCESSO DE OTIMIZAÇÃO DOS PESOS ..	97
QUADRO 7.2	: PESOS OBTIDOS PELOS MÉTODOS: LP, BFGS E REGIÃO DE CONFIANÇA .....	102

## LISTA DE SIGLAS

BFGS	: Broyden, Fletcher, Goldfarb e Shanno
GPS	: <i>Global Positioning System</i>
IAG	: <i>International Association of Geodesy</i>
IME	: Instituto Militar de Engenharia
PVPI	: Problema de Valor Próprio Inverso

## LISTA DE SÍMBOLOS

- $m_c$  : representação de um modelo quadrático obtido utilizando a expansão truncada de ordem 2 de Taylor;
- $M_k$  : modelo quadrático obtido para aplicação do método de região de confiança
- $\mathbf{x}^+$  : representa o vetor minimizador de um modelo quadrático ( não confundir com a pseudo-inversa de um vetor);
- $\mathbf{x}^*$  : representa um vetor minimizador (local ou global) de uma função qualquer;
- $\lambda_*$  : representação de um escalar (utilizado no método de região de confiança);
- $\lambda^*$  : valor próprio pretendido em um processo de otimização dos pesos;
- $\mu_i$  : i-ésimo valor próprio associado a matriz de covariâncias dos parâmetros estimados em um processo de otimização dos pesos
- $\mu$  : escalar utilizado no método de Newton com busca linear o qual garante a positividade da matriz hessiana;
- $\sigma$  : representa uma permutação de valores próprios associado ao método LP (*lift and projection*)
- $\Lambda^*$  : matriz diagonal que apresenta em sua diagonal principal um conjunto de valores próprios pré-definidos;
- 1 : indicação de inversa ordinária quando sobrescrito no símbolo da matriz;
- + : indicação da inversa generalizada de Moore-Penrose ou pseudo-inversa quando sobrescrito no símbolo da matriz ou vetor (exceto para  $\mathbf{x}^+$ );
- $\|(\bullet)\|$  : norma euclidiana de um vetor qualquer
- $\sum_{i=1}^p$  : indica somatório desde i até p
- $\prod_{i=1}^p$  : indica produtório desde i até p
- ! : deve ser
- =

## RESUMO

Neste trabalho apresenta-se o problema de segunda ordem em redes geodésicas que pode, em geral, ser solucionado por uma matriz critério ou por um problema de valor próprio inverso. Os métodos baseados em matriz critério são revistos e exemplificados. Apresentam-se ainda os fundamentos matemáticos sobre a otimização matemática, os quais possibilitam a aplicação de métodos globalizados na solução de um problema não linear. O problema de valor próprio inverso é solucionado globalizando o método de Newton usando as estratégias busca linear e região de confiança, método quase-Newton com a fórmula BFGS. O método LP (*lift and projection*) é apresentado para solução de um problema de valor próprio inverso. Os algoritmos são aplicados na otimização de rede geodésicas bidimensionais, projetadas para atender a uma precisão pré-definida. Os resultados são analisados e discutidos.

## ABSTRACT

In this work the problem of second order design in geodetic network is presented that can, in general, to be solved by a criterion matrix or an inverse eigenvalue problem. The methods based on criterion matrix are reviewed and exemplify. The fundamentals mathematical on the mathematical optimization are presented, which make possible the application of global methods in the solution of a nonlinear problem. The inverse eigenvalue problem is solved by global Newton's method using the strategies line search and trust region, quasi-Newton method with BFGS formula. The method LP (lift and projection) is presented for solution of a inverse eigenvalue problem. The algorithms are applied in the optimization of a bidimensional geodetic network when it is specified the final precision. The results are analyzed and argued.

## 1 INTRODUÇÃO

A Geodésia é a ciência que tem por objetivo determinar a forma e as dimensões da Terra e os parâmetros definidores do campo de gravidade (GEMAEL, 1994, p. 15).

Qualquer medida tomada apresenta erros. Estes erros são provenientes de causas como, falhas humanas, efeitos físicos e equipamentos imperfeitos. Para minimizar tais erros repete-se a mensuração várias vezes. Com a pluralidade e discrepância deste conjunto de observações obriga-se a extrair um resultado único e que apresente maior confiança na grandeza medida. Isto é possível por meio de técnicas que levam as observações obedecerem a um modelo matemático (ajustamento de observações), o qual permite a obtenção de um resultado único e com alto grau de confiabilidade.

Muitas vezes as observações não apresentam o mesmo grau de confiança, então, podem-se atribuir pesos diferenciados valorizando as que possuem melhor qualidade. Para tanto a atribuição de pesos às observações pressupõe o conhecimento da precisão de cada medida.

O problema de otimização de pesos é conhecido na literatura geodésica como planejamento ou projeto de segunda ordem no qual se conhece a matriz planejamento  $A$ , também conhecida como matriz dos coeficientes ou matriz de configuração e os pesos são tratados como incógnitas.

A idéia básica da otimização de uma rede geodésica está no reconhecimento que é possível, fornecida a sua configuração, estimar a qualidade da rede antes de qualquer medição ser realizada. Esta idéia abre a possibilidade de analisar detalhadamente o projeto com respeito a sua precisão, confiabilidade, custos e ainda melhorar partes da rede, se necessário.

A partir daí pode-se obter pesos otimizados para as observações fazendo-se uso de modelos “ótimos” para as matrizes de covariâncias dos parâmetros estimados matrizes estas chamadas de matriz critério. Desta forma qualquer matriz positiva definida ou semi-definida positiva pode ser usada como uma matriz critério desde que represente uma situação ideal para a precisão, uma vez que não há necessidade que a matriz de covariâncias efetivamente obtida na prática seja igual à matriz critério. Um

caso particular de uma matriz critério homogênea e isotrópica é a matriz identidade, isto é,  $\mathbf{Q}_x = \mathbf{I}$

Uma outra possibilidade de otimização de pesos é o uso de valores próprios. Tendo em vista que o uso dos valores próprios e vetores próprios é um diagnóstico da qualidade de um ajustamento, pode-se formular a otimização de pesos das observações geodésicas por meio de condições impostas aos valores próprios da matriz de covariâncias, ou seja, os valores próprios da matriz de covariâncias apresentam-se como função objetivo no processo de otimização. Em suma pretende-se que a rede apresente valores próprios pré-estabelecidos.

Esta situação tem resolução através da solução iterativa de um Problema de Valor Próprio Inverso aplicado sobre a matriz dos coeficientes das equações normais.

Alguns métodos para solução do problema de valor próprio inverso, baseados no método de Newton clássico podem ser encontrados em FRIEDLAND et al. (1987, p. 639-643).

Neste trabalho outras técnicas para solução do problema de valor próprio inverso, especificamente na otimização dos pesos são analisadas. Modela-se o problema por meio de uma função de minimização para a qual serão utilizados, iterativamente, modelos quadráticos explorando a possibilidade da utilização de métodos que melhorem as garantias de solubilidade do problema.

Os métodos utilizados para solução da função de minimização são: o Método de Newton no qual se utilizou as estratégias de globalização do método por busca linear e região de confiança, o método quase-Newton globalizado BFGS que utiliza modelos definido-positivo para a matriz de segunda derivadas ou matriz hessiana e ainda o método LP (*lift and projection*) que não faz uso de derivadas primeiras e nem de derivadas segundas. Cada método foi aplicado na otimização dos pesos com base em valor próprio inverso variando-se a configuração da rede e ainda o número de valores próprios pré-definidos que a matriz de covariâncias da rede deve apresentar. Os resultados foram analisados e discutidos.

## 1.1 CONSIDERAÇÕES SOBRE PROJETOS DE REDES GEODÉSICAS

Uma rede geodésica é constituída por um conjunto de pontos materializados no terreno, com suas posições definidas em relação a um sistema geodésico de referência. Para sua determinação são utilizadas medições geodésicas que podem ser observações de ângulos, direções, distâncias, diferenças de níveis associadas a observações gravimétricas ainda por técnicas astronômica ou espacial, por exemplo, observações GPS. A importância de se planejar ou projetar uma rede geodésica, para posterior execução, está no fato que pretende-se para o levantamento geodésico custos reduzidos e a obtenção de posições precisas após o ajustamento da rede. Com a finalidade de alcançar o objetivo de precisão e custo reduzido, em geral, a experiência é a linha de conduta de um geodesta na concepção da rede geodésica.

O projeto da rede é concluído após um reconhecimento no campo, onde a escolha dos pontos pode ser mantida ou alterada por problemas de visibilidade, rigidez geométrica, acesso entre outros. Após o projeto concluído segue-se a campanha de medição onde observações (variável aleatória) são obtidas com a instrumentação disponível. De posse dos valores observados as posições (parâmetros) podem ser obtidas por meio de um ajustamento utilizando-se para isso, por exemplo, o Método dos Mínimos Quadrados.

O sucesso na implantação da rede depende da qualidade final da mesma, que deve ser ótimo sob algum critério estabelecido em função da qualidade do projeto.

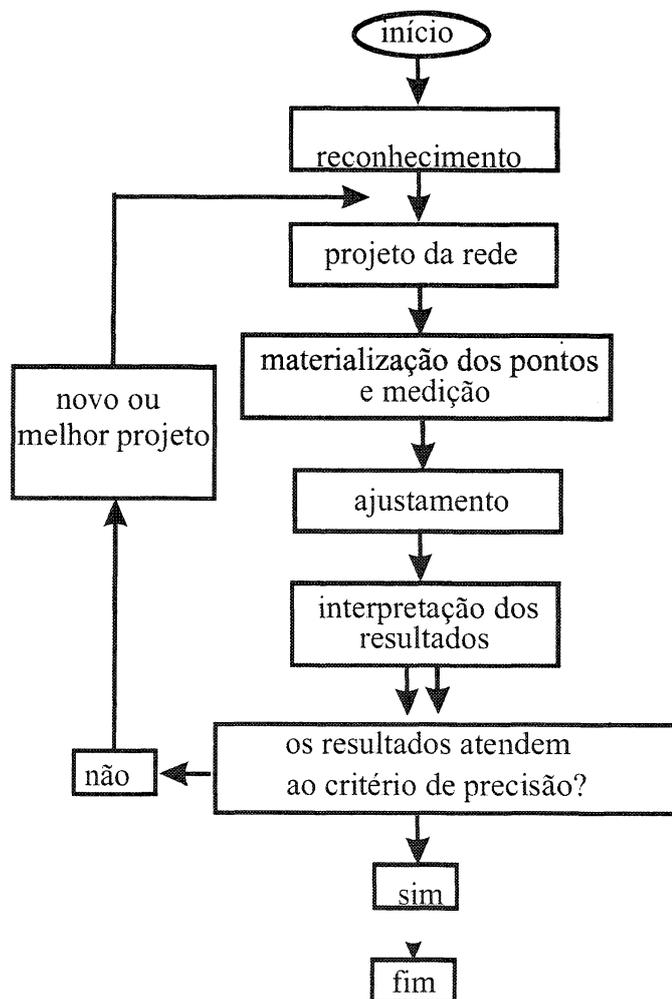
Após a obtenção das posições, por meio de um ajustamento, é necessário analisar a qualidade do levantamento, com o intuito de possibilitar a detecção de deficiências (teste da variância e análise da precisão) e fragilidades (rigidez geométrica, número de redundância e erros grosseiros) na rede.

Nesta concepção deficiências e/ou fragilidades se forem detectadas, serão corrigidas ou ao menos minimizados os seus efeitos através de um novo levantamento em campo onde novos pontos serão escolhidos, outros equipamentos utilizados, outras medições ou remedições realizadas, um novo ajustamento deverá ser realizado e conseqüentemente uma nova análise de qualidade deve ser feita. Todo este processo

representa um gasto adicional, tanto do ponto de vista financeiro quanto de tempo e mesmo assim este novo trabalho pode não ter a precisão esperada.

A figura 1.1 apresenta o fluxograma no estabelecimento de uma rede como descrito acima.

FIGURA 1.1 - ESTABELECIMENTO DE UMA REDE GEODÉSICA



Posteriormente a este trabalho algumas indagações são pertinentes ao levantamento:

- O plano de observações (equipamentos, disposição dos pontos, tipo de observação) foi adequado?
- A ocupação dos pontos foi realizada de forma a minimizar custos (tempo e financeiro)?

Para minimizar as deficiências que eventualmente podem ocorrer, tanto durante quanto posterior a um levantamento, muitas vezes a prática é a única ferramenta usada para melhorar a concepção do levantamento.

A experiência é um dado que não deve ser descartado em um levantamento geodésico, porém, métodos analíticos (ou não) juntamente com critérios de precisão, confiabilidade (rigidez) e custos são ferramentas matemáticas que podem ser utilizadas concomitantemente com a experiência, como linhas de conduta no projeto geodésico.

Melhorar o conhecimento sobre a rede na fase de planejamento, minimizar fragilidades e corrigir deficiências, antes de qualquer campanha de medição é parte da Otimização do Projeto de Redes Geodésicas, que fornece as informações a cerca do trabalho a ser realizado, de tal forma que torna possível projetar e conceber melhor o levantamento. A otimização do projeto de redes geodésicas também é denominada pré-análise, uma vez que é realizada sem que se efetue qualquer medição nem o conseqüente ajustamento (SÁ, 1985, p. 9).

O termo otimização vem sendo utilizado, pela inserção de funções que representam precisão, confiabilidade e custo e ainda métodos computacionais de otimização que são ferramentas matemáticas utilizadas para maximizar ou minimizar funções sujeitas ou não a restrições nas variáveis.

De forma geral pode-se representar uma função de otimização de redes geodésicas como,

$$\mathbf{F}(\mathbf{x}, \mathbf{P}) = p(\mathbf{x}, \mathbf{P}) + r(\mathbf{x}, \mathbf{P}) + c(\mathbf{x}, \mathbf{P}) \quad (1.1)$$

onde  $\mathbf{F}(\mathbf{x}, \mathbf{P})$  representa a função a ser minimizada ou maximizada e esta está em função da posição dos pontos e/ou da precisão de cada observação com a contribuição  $p(\mathbf{x}, \mathbf{P})$ ,  $r(\mathbf{x}, \mathbf{P})$  e  $c(\mathbf{x}, \mathbf{P})$  da precisão, confiabilidade e custo, respectivamente.

Tradicionalmente, os problemas de otimização são classificados em 4 ordens, sendo que cada ordem é classificada de acordo com a grandeza que se quer obter no processo de otimização. Desta forma GRAFAREND (1974) classifica os projetos de otimização nas seguintes ordens:

- Projeto de Zero Ordem: trata de escolher um sistema geodésico de referência ótimo;

- Projeto de Primeira Ordem: trata de projetar uma configuração ótima para uma rede;
- Projeto de Segunda Ordem: seleção dos pesos para as observações que servirão de referência na realização de uma rede geodésica (pesos otimizados);
- Projeto de Terceira Ordem: inclusão de novas observações para melhorar uma rede já existente.

Quando os projetos de Primeira e Segunda ordem são resolvidos simultaneamente, é chamado de Projeto Combinado.

Basicamente os procedimentos que podem ser usados para resolver os problemas de otimização nas suas diferentes ordens são: o método de tentativa e erro e o método analítico.

O método de tentativa e erro é um método que visa projetar uma rede usando recursos de simulação em computador que incorpora técnicas empíricas e norteadas pela experiência do projetista. Nesta técnica postula-se uma solução envolvendo critérios de custo e precisão. Algumas redes são simuladas até que os critérios sejam satisfeitos. Se ocorrer que nenhum dos critérios for satisfeito um novo plano é postulado (por exemplo, alterando o projeto original) e a rede é novamente simulada. O procedimento é repetido até que uma rede satisfatória seja encontrada, porém isto não quer dizer que esta rede seja a ótima. Em resumo, para uma dada rede que deve ser projetada por tentativa e erro se obedece os seguintes passos:

1. Especificar critérios para precisão e confiabilidade;
2. Selecionar um plano de observações (pontos no terreno, tipos de observações e pesos);
3. Calcular a matriz de covariâncias dos parâmetros estimados e deduzir as quantidades que representam a precisão e a confiabilidade da rede;
4. Se os valores obtidos no passo 3 são satisfatórios com os estabelecidos no passo 1, então vai-se para o passo 5, caso contrário altera-se o plano de observações (removendo ou adicionando observações ou diminuindo ou aumentando os pesos das observações) e retorna-se para o passo 3;

5. Obter o custo da rede e considerar a possibilidade de retornar ao passo 2 e reiniciar com um tipo diferente de rede. Parar quando se acredita que a rede ótima foi encontrada.

Um conjunto de precisões pré-definidas, as quais são idealizadas em um sentido ótimo para cada ponto da rede, em geral são caracterizadas por elipses de erros degeneradas em círculos idênticos (características de homogeneidade e isotropismo).

Os cálculos de um ajustamento simulado com as grandezas dadas nos passos (1), (2) e (3) são realizados e definem-se elipses de erro para as posições da rede (parâmetros). Os resultados são comparados com as precisões pré-definidas e executa-se o passo (4). Caso as elipses de erro calculadas excedam as regiões limitadas pelos círculos de erro postulados, uma nova simulação será feita refinando-se a simulação anterior, seja entrando com novas observações ou subtraindo observações e/ou “melhorando” as precisões das observações anteriores. O procedimento é repetido até que uma rede satisfatória seja encontrada. As grandezas obtidas no projeto final devem ser realizáveis no levantamento.

Os dados de entrada tendem a serem próximos do ideal quando o projetista usa sua experiência e habilidade. Neste sentido, o método de tentativa e erro possibilita que o projetista use sua experiência para melhor simular a rede.

Em contraste, os métodos analíticos oferecem algoritmos específicos para a solução de cada projeto os quais não requerem intervenção humana. O termo analítico é usado para descrever um método que resolve um projeto particular pela aplicação de algum algoritmo matemático. A aplicação de tal algoritmo conduzirá automaticamente à solução de uma rede que irá satisfazer os requerimentos de qualidade pré-definidos e esta rede será ótima em algum sentido matemático. Posteriormente métodos analíticos serão aplicados na otimização de redes geodésicas.

## 1.2 HISTÓRICO

O ponto inicial das técnicas de otimização analítica, objetivando medições geodésicas é devido à dissertação de F. R. Helmert de 1868, intitulada “Studien über

rationelle Vermessungen im Gebiet der höheren Geodäsie”. Neste trabalho é proposto um modelo para racionalizar um levantamento, tentando encontrar regras para a localização de pontos de uma rede em função do tipo de medição e do número de observações. Helmert estabeleceu postulados a respeito de um projeto ótimo (*optimal design*) que ainda hoje são aceitos. Como exemplo utilizou o valor mínimo do traço da matriz de covariâncias das coordenadas dos pontos estimados como critério de precisão e impôs ao planejamento restrições de custo. Nesta mesma linha outros geodestas fizeram contribuições como SCHEREIBER (1882), JUNG (1924), WOLF (1961). Todos buscavam minimizar alguma função objetivo, que descreve custo, precisão ou confiabilidade em um projeto geodésico. Baarda em 1962 propôs uma metodologia diferente, as chamadas matrizes critério as quais são uma melhor aproximação para a matriz de covariâncias dos parâmetros estimados. Estas matrizes critério possuem uma estrutura ideal que é especificada em cada caso. No caso de redes geodésicas bi ou tri-dimensionais a estrutura ideal pode ser estabelecida pela estrutura de Taylor-Karman que foram introduzidas nas Ciências Geodésicas por Erik W. Grafarend em 1972. Em GRAFAREND e SCHAFFRIN (1979) matrizes com estrutura de Taylor-Karman foram apresentadas para observações de azimutes, ângulos e distâncias. Neste período é constituído pela Associação Internacional de Geodésia (IAG) um grupo especial de estudo neste campo sob o título “Otimização do Projeto de Redes Geodésicas”.

Nesta época muitos estudos eram devotados ao projeto de segunda ordem, porém aos aspectos de confiabilidade ainda não eram dadas as devidas atenções. Nesta concepção Van Mierlo em 1981 apresenta o trabalho “*Second Order Design: precision and reliability aspects*” onde apresentou estudos sobre a associação entre a precisão e a confiabilidade quando o intuito é projetar uma rede. Em suma os pesos obtidos para a rede projetada deve atender a critérios de precisão para as parâmetros e confiabilidade para observações.

No Brasil em 1985 Carlos César Paiva de Sá apresenta a dissertação de mestrado intitulada Otimização de Observações em Redes Horizontais pelo IME no qual utilizou matrizes critério de covariância com estrutura Taylor-Karman para solucionar o projeto de segunda ordem em redes geodésicas horizontais.

O uso de valores próprios, de forma direta, em projetos de redes geodésicas tem seu ponto inicial no trabalho de Rene Jäger em 1988 intitulado “Analyse und Optimierung geodätischer Netze nach spektralen Kriterien und mechanische Analogien” no qual apresenta uma formulação matemática que possibilita extrair pesos se é estabelecido um conjunto de valores próprios para a matriz de covariâncias dos parâmetros estimados. Hubert Kaltenbach em 1992 apresenta a otimização da configuração da rede (Projeto de Primeira Ordem) baseado em um problema de valor próprio inverso onde soluciona um sistema de equações não-lineares utilizando o método de Newton clássico.

Neste mesmo período DEREN e YONGQIAN (1991) apresentam um método que possibilita resolver problemas de otimização, de primeira e segunda ordem, construindo uma matriz critério de confiabilidade obtida de forma empírica.

Vários pesquisadores vem contribuindo significativamente para a evolução e aplicabilidade da otimização em projetos de redes geodésicas. Em tempos recentes os problemas de otimização são tratados nos trabalhos de AMIRI-SIMKOOEI (2004), SCHMITT (1997), SCHWIEGER (2001) e STOPAR (2001) e entre outros, os quais se ocuparam com os problemas de projetar redes geodésicas.

### 1.3 COOPERAÇÃO E SUPORTE PARA SUA REALIZAÇÃO

- a) Curso de Pós-Graduação em Ciências Geodésicas, do Departamento de Geomática, Setor Ciências da Terra, da Universidade Federal do Paraná, mediante a linha de pesquisa em geodésia, sub-área de Otimização de Levantamentos geodésicos;
- b) Orientação do Prof. Dr. Quintino Dalmolin.

### 1.4 OBJETIVOS

#### 1.4.1 Objetivo Geral

Implementar um método para otimização dos pesos considerando um problema de valor próprio inverso com considerações sobre algoritmos de otimização matemática e garantias quanto à solubilidade de um projeto de rede geodésica.

#### 1.4.2 Objetivos Específicos

- a) rever os modelos de otimização dos pesos;
- b) estabelecer diferenças entre as metodologias de otimização dos pesos baseadas em matriz critério e um problema de valor próprio inverso;
- c) apresentar as vantagens e desvantagens sobre as formulações de um problema de valor próprio inverso;
- d) apresentar os fundamentos da modelagem e solubilidade de um problema não linear;
- e) estabelecer critérios que possibilitem a otimização baseada na metodologia de um problema de valor próprio inverso;
- f) modelar o problema de otimização dos pesos através de uma função de minimização;
- g) apresentar os algoritmos de otimização matemática que possibilitam a solução de um problema de valor próprio inverso e em consequência a obtenção de uma rede geodésica otimizada;
- h) aplicar os algoritmos de otimização em problema de otimização dos pesos das observações modelados por um problema de valor próprio inverso;
- i) apresentar resultados da aplicabilidade dos algoritmos de otimização matemáticas em problemas de otimização de redes geodésicas.

#### 1.5 CONTRIBUIÇÃO E ESTRUTURAÇÃO DA PESQUISA

Este trabalho visa contribuir com uma metodologia para a otimização dos pesos em levantamentos geodésicos ou, em termos gerais, no estabelecimento de uma

rede geodésica, quando esta necessita ser projetada para que atenda a uma precisão pré- definida. A metodologia utilizada é baseada em valores próprios cuja modelagem é solucionada com um problema de valor próprio inverso.

As garantias que o projeto da rede obterá êxito e a rede poderá ser otimizada são ampliadas com a aplicação de algoritmos de otimização matemática sobre uma função de minimização construída com o problema de valor próprio inverso.

Pretende-se contribuir ainda com uma revisão dos modelos aplicados à otimização dos pesos, anteriores à proposta aqui apresentada, facilitando assim uma comparação e em conseqüência o estabelecimento das diferenças (vantagens e/ou desvantagens) entres as metodologias existentes neste campo de estudo.

Espera-se ainda que este trabalho contribua para facilitar o estudo e a aplicabilidade deste tema em levantamentos geodésicos que necessitem serem projetados para uma melhor concepção do seu ajustamento e qualidade final do empreendimento.

A pesquisa esta estruturada em 8 seções:

A primeira seção, seção introdutória, expõe os fundamentos clássicos e o estado da arte do problema de otimização de redes geodésicas.

Na segunda seção estão apresentados os conceitos fundamentais que possibilitam a modelagem e a solução dos problemas de otimização dos pesos.

Na terceira seção são apresentados os fundamentos que permitem modelar o problema de otimização dos pesos das observações geodésicas quando a metodologia é baseada em matriz critério.

Na quarta seção são estabelecidos os fundamentos que permitem modelar o mesmo problema de otimização dos pesos por meio do conceito de valor próprio.

Na quinta seção é apresentada a modelagem do problema de otimização dos pesos baseada em um valor próprio, ou seja, é apresentado o problema de valor próprio inverso que permite a solução do problema exposto na seção 4.

Na sexta seção são apresentados os algoritmos de otimização matemática bem como seus fundamentos, os quais são ferramentas úteis na solução de um problema de valor próprio inverso e em conseqüência no estabelecimento de uma rede projetada.

Na sétima seção são apresentados três experimentos utilizando o processo de otimização com a metodologia proposta neste trabalho cuja solução é procurada com o método de Newton com estratégia de busca linear e região de confiança, método quase-Newton BFGS e ainda o método LP (*lift and projection*).

Na oitava seção o trabalho é concluído com a análise dos resultados.

## 2 CONCEITOS FUNDAMENTAIS

Neste capítulo apresentam-se as ferramentas matemáticas que serão úteis na formulação e solução dos problemas de otimização dos pesos. Os conceitos aqui abordados fundamentam e modelam os processos de otimização.

Apresenta-se a definição de uma matriz ortogonal, da inversa ordinária de uma matriz quadrada e inversa generalizada de Moore-Penrose, que se aplica também a uma matriz retangular e possibilita a solução de um sistema de equações lineares no sentido do método dos mínimos quadrados com solução de norma mínima. Um algoritmo para construção da pseudo-inversa de uma matriz é apresentado uma vez que esta é o fundamento para obtenção dos pesos com respeito a matriz critério de covariâncias.

Em particular os produtos especiais de matrizes que serão abordados neste capítulo, produtos de Kronecker e de Khatri-Rao conjuntamente com o operador vec possibilitam transformar um produto de matrizes em um sistema de equações lineares, conduzindo à modelagem do problema de otimização dos pesos fundamentado em matrizes de covariâncias pré-definidas.

Por fim a definição de um problema de valor próprio e vetores próprios, especial e generalizado, são apresentados. O conceito de valor próprio é o fundamento para a otimização dos pesos das observações geodésicas quando se emprega a metodologia de um problema de valor próprio inverso como sugestionado neste trabalho.

### 2.1 MATRIZ ORTOGONAL

Uma matriz real  $\mathbf{A}$  é ortogonal se for válida a relação,

$$\mathbf{A}^t = \mathbf{A}^{-1} \quad (2.1)$$

ou de forma equivalente

$$\mathbf{A}\mathbf{A}^t = \mathbf{A}^t\mathbf{A} = \mathbf{I}. \quad (2.2)$$

Uma matriz ortogonal é necessariamente quadrada e inversível, tendo em vista a definição (2.1).

A condição  $\mathbf{A}^t \mathbf{A} = \mathbf{I}$  mostra que as colunas de  $\mathbf{A}$  formam um conjunto ortonormal de vetores.

## 2.2 INVERSA ORDINÁRIA

Uma matriz quadrada  $\mathbf{A}$  de dimensão  $m \times m$  é não-singular ou regular ou inversível se existe uma única matriz  $\mathbf{B}$  de dimensão  $m \times m$  tal que  $\mathbf{AB} = \mathbf{I}$ . A matriz  $\mathbf{B}$  é denotada por  $\mathbf{A}^{-1}$ , e diz-se que  $\mathbf{B}$  é a inversa ordinária de  $\mathbf{A}$  e vale a relação:

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \quad (2.3)$$

A partir do determinante da matriz  $\mathbf{A}$  e da matriz dos cofatores de  $\mathbf{A}$ , obtém-se sua inversa pela relação:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \mathbf{A}^*, \quad (2.4)$$

onde  $\mathbf{A}^*$  denota a matriz cofatora de  $\mathbf{A}$ .

Observa-se a partir de (2.3) que se  $\det(\mathbf{A}) = 0$ ,  $\mathbf{A}$  não possui inversa ordinária.

## 2.3 INVERSA GENERALIZADA E INVERSA GENERALIZADA DE MOORE-PENROSE

Seja  $\mathbf{A}$  uma matriz de dimensão  $m \times n$ . Qualquer matriz  $\mathbf{B}$  satisfazendo  $\mathbf{ABA} = \mathbf{A}$  é uma inversa generalizada de  $\mathbf{A}$ . Em geral uma inversa generalizada não é única.

Uma matriz  $\mathbf{B}$  é chamada inversa generalizada de Moore-Penrose de  $\mathbf{A}$  ou pseudo-inversa de  $\mathbf{A}$ , se satisfaz as quatro condições seguintes (RAO e MITRA, 1971, p. vii):

$$\mathbf{ABA} = \mathbf{A} , \quad (2.5)$$

$$\mathbf{BAB} = \mathbf{B} , \quad (2.6)$$

$$(\mathbf{AB})^t = \mathbf{AB} , \quad (2.7)$$

$$(\mathbf{BA})^t = \mathbf{BA} . \quad (2.8)$$

A inversa generalizada de Moore-Penrose ou pseudo-inversa de  $\mathbf{A}$ , denotada pelo símbolo  $\mathbf{A}^+$ , existe para qualquer matriz  $m \times n$  e é única.

Propriedades da pseudo-inversa:

$$\mathbf{A}^+ \text{ é única} \quad (2.9)$$

$$\mathbf{A}^+ = \mathbf{A}^{-1} , \text{ se } \mathbf{A} \text{ é não - singular,} \quad (2.10)$$

$$(\mathbf{A}^+)^+ = \mathbf{A} ; \quad (2.11)$$

$$(\mathbf{A}^t)^+ = (\mathbf{A}^+)^t ; \quad (2.12)$$

$$\mathbf{A}^t \mathbf{A} \mathbf{A}^+ = \mathbf{A}^+ \mathbf{A} \mathbf{A}^t = \mathbf{A}^t ; \quad (2.13)$$

$$\mathbf{A}^t (\mathbf{A}^+)^t \mathbf{A}^+ = \mathbf{A}^+ (\mathbf{A}^+)^t \mathbf{A}^t = \mathbf{A}^+ ; \quad (2.14)$$

$$(\mathbf{A}^t \mathbf{A})^+ = \mathbf{A}^+ (\mathbf{A}^+)^t ; \quad (2.15)$$

$$\mathbf{A}^+ = (\mathbf{A}^t \mathbf{A})^+ \mathbf{A}^t = \mathbf{A}^t (\mathbf{A} \mathbf{A}^t)^+ ; \quad (2.16)$$

$$\mathbf{a}^+ = \frac{\mathbf{a}^t}{\|\mathbf{a}\|_2^2} = \frac{\mathbf{a}^t}{\mathbf{a}^t \mathbf{a}} , \text{ se } \mathbf{a} \text{ é um vetor não nulo.} \quad (2.17)$$

Nas Ciências Geodésicas a pseudo-inversa é usada em redes livres, ou seja, uma rede que possa ser ajustada sem a fixação de nenhuma injunção inicial. A rede livre produz singularidade na matriz das equações normais, isto é, não admite a inversa ordinária.

Um algoritmo numericamente estável baseado na decomposição de valor singular, para obter a inversa generalizada de Moore-Penrose ou pseudo-inversa de uma matriz qualquer  $\mathbf{A}$  é descrito pelas seguintes considerações:

Para qualquer matriz  $\mathbf{A}$ , não necessariamente quadrada, o produto  $\mathbf{A}^t \mathbf{A}$  apresenta valores próprios não negativos. A raiz quadrada positiva de cada um destes

valores próprios é chamado de valor singular de  $\mathbf{A}$ . Além disso, existem matrizes unitárias  $\mathbf{U}$  e  $\mathbf{V}$  tais que é válida a relação,

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{V}^t \quad (2.18)$$

onde  $\mathbf{D}$  é uma matriz diagonal tendo na diagonal principal todos os valores singulares positivos da matriz  $\mathbf{A}$  e  $\mathbf{0}$  é uma matriz onde todos seus elementos são zero. A matriz bloco,

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (2.19)$$

é do mesmo tipo da matriz  $\mathbf{A}$  e, em consequência, é quadrada somente quando  $\mathbf{A}$  é quadrada.

A decomposição (2.18) gera uma fórmula numericamente estável para obtenção da pseudo-inversa como,

$$\mathbf{A}^+ = \mathbf{V} \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^t \quad (2.20)$$

a qual pode ser simplificada para

$$\mathbf{A} = \mathbf{V}_1 \mathbf{D}^{-1} \mathbf{U}_1 \quad (2.21)$$

onde as colunas de  $\mathbf{V}_1$  são os vetores próprios que correspondem aos valores próprios positivos e  $\mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \mathbf{D}^{-1}$ .

## 2.4 SOLUÇÃO DE UM SISTEMA DE EQUAÇÕES LINEARES PELOS MÍNIMOS QUADRADOS

A pseudo-inversa, do ponto de vista prático, responde a uma pergunta bastante natural que ocorre em diferentes aplicações. Se é impossível encontrar  $\mathbf{x}$  tal que  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , então quais são os vetores  $\mathbf{x}$  tais que o erro  $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|$  é o menor possível e qual entre esses vetores  $\mathbf{x}$  tem a menor norma.

Uma solução pelos mínimos quadrados de um sistema de equações lineares  $\mathbf{Ax} = \mathbf{b}$  é o vetor com menor norma euclidiana que minimiza  $\|\mathbf{Ax} - \mathbf{b}\|_2$ . Esse vetor é  $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$ . (2.22)

Quando  $\mathbf{A}$  possui inversa (2.22) reduz-se à  $\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$ , que é a única solução. Para sistemas possíveis, que admitem infinitas soluções (2.22) identifica a solução que tem a norma euclidiana mínima. A (2.22) identifica, também, uma solução para sistemas impossíveis, a melhor em termos de mínimos quadrados.

Para solução do método dos mínimos quadrados linear e não-linear consultar, por exemplo, GARNÉS (2001).

## 2.5 PRODUTO DE KRONECKER

Seja  $\mathbf{A} = (a_{i,j})$  e  $\mathbf{B} = (b_{i,j})$  matrizes quaisquer de dimensão  $m \times n$  e  $p \times q$  respectivamente. Então, a matriz  $\mathbf{C} = (a_{ij} \mathbf{B})$  é uma matriz de dimensão  $mp \times nq$ . A matriz resultante  $\mathbf{C}$  é chamada de produto de Kronecker entre a matriz  $\mathbf{A}$  e  $\mathbf{B}$ , e é usualmente representado por  $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$ , isto é,

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11} \mathbf{B} & a_{12} \mathbf{B} & \cdots & a_{1n} \mathbf{B} \\ a_{21} \mathbf{B} & a_{22} \mathbf{B} & \cdots & a_{2n} \mathbf{B} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} \mathbf{B} & a_{m2} \mathbf{B} & \cdots & a_{mn} \mathbf{B} \end{bmatrix}. \quad (2.23)$$

Exemplo

Dadas as matrizes,  $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$  e  $\mathbf{B} = \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}$  o produto de Kronecker

entre  $\mathbf{A}$  e  $\mathbf{B}$  é obtido seguindo a definição (2.23) como segue,

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} 1 \times \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix} & 2 \times \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix} \\ 3 \times \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix} & 4 \times \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix} \end{bmatrix} \text{ como resultado tem-se}$$

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} 5 & 6 & 7 & 10 & 12 & 14 \\ 8 & 9 & 10 & 16 & 18 & 20 \\ 15 & 18 & 21 & 20 & 24 & 28 \\ 24 & 27 & 30 & 32 & 36 & 40 \end{bmatrix}.$$

## 2.6 PRODUTO DE KHATRI-RAO

Este tipo de produto entre matrizes foi estudado por Khatri e Rao em 1968 e sua forma geral é expressa como se segue.

$$\text{Sejam } \mathbf{A} = [\mathbf{A}_1 \quad \cdots \quad \mathbf{A}_i \quad \cdots \quad \mathbf{A}_k] \text{ e } \mathbf{B} = [\mathbf{B}_1 \quad \cdots \quad \mathbf{B}_i \quad \cdots \quad \mathbf{B}_k]$$

duas matrizes particionadas nas suas colunas e com o mesmo número de partições.

Define-se o produto de Khatri-Rao como

$$\mathbf{A} \odot \mathbf{B} = [\mathbf{A}_1 \otimes \mathbf{B}_1 \quad \cdots \quad \mathbf{A}_i \otimes \mathbf{B}_i \quad \cdots \quad \mathbf{A}_k \otimes \mathbf{B}_k] \quad (2.24)$$

onde  $\mathbf{A}_i \otimes \mathbf{B}_i$  é o produto de Kronecker entre as  $i$ -ésimas partições  $\mathbf{A}_i$  e  $\mathbf{B}_i$ . O produto de Khatri-Rao pode ser estabelecido para o caso particular onde cada partição é um vetor coluna das matrizes  $\mathbf{A}$  e  $\mathbf{B}$  respectivamente.

Propriedades do produto de Khatri-Rao

$$1) (\mathbf{A} \odot \mathbf{B}) \odot \mathbf{C} = \mathbf{A} \odot (\mathbf{B} \odot \mathbf{C}) \quad (2.25)$$

$$2) (\mathbf{C} \otimes \mathbf{D})(\mathbf{A} \odot \mathbf{B}) = \mathbf{CA} \odot \mathbf{DB} \quad (2.26)$$

Exemplo

$$\text{Dadas as matrizes, } \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ e } \mathbf{B} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \text{ ambas com o mesmo número de}$$

partições (no caso colunas). O produto de Khatri-Rao dado pela definição (2.24) é obtido como se segue,

$$\mathbf{A} \odot \mathbf{B} = \left[ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \otimes \begin{bmatrix} 5 \\ 7 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 4 \end{bmatrix} \otimes \begin{bmatrix} 6 \\ 8 \end{bmatrix} \right] \text{ que usando a definição (2.23) conduz a}$$

$$\mathbf{A} \odot \mathbf{B} = \left[ \begin{array}{cc} 1 \times \begin{bmatrix} 5 \\ 7 \end{bmatrix} & 2 \times \begin{bmatrix} 6 \\ 8 \end{bmatrix} \\ 3 \times \begin{bmatrix} 5 \\ 7 \end{bmatrix} & 4 \times \begin{bmatrix} 6 \\ 8 \end{bmatrix} \end{array} \right] \text{ cujo resultado é}$$

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} 5 & 12 \\ 7 & 16 \\ 15 & 24 \\ 21 & 32 \end{bmatrix}.$$

## 2.7 OPERADORES $\text{vec}$ e $\text{vecd}$

Em muitos casos, se faz necessário expressar uma matriz  $\mathbf{B}$  na forma de vetor utilizando seus vetores coluna. Desta forma, seja  $\mathbf{B}$  uma matriz  $m \times n$  expressa como  $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_i, \dots, \mathbf{b}_n)$ , onde  $\mathbf{b}_i = (a_{1i}, \dots, a_{mi})$  são vetores de dimensão  $m \times 1$  com  $i = 1, \dots, n$ . Se todas as colunas forem colocadas uma sobre as outras (empilhamento das colunas) obtém-se um vetor de dimensão  $mn \times 1$  da seguinte forma (HENDERSON; SEARLE, 1979, p. 65):

$$\text{vec}(\mathbf{B}) = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_i \\ \vdots \\ \mathbf{b}_n \end{bmatrix} \quad (2.27)$$

onde,  $\text{vec}(\mathbf{B})$  é algumas vezes referido como a representação vetorial da matriz  $\mathbf{B}$  (WANG; CHOW, 1994, p.54).

O operador  $\text{vecd}$  estabelece a representação vetorial de uma matriz diagonal. Então seja  $\mathbf{B}$  uma matriz diagonal de dimensão  $m \times m$  como,

$$\mathbf{B} = \begin{bmatrix} b_{11} & & & & \\ & \ddots & & & \\ & & b_{ii} & & \\ & & & \ddots & \\ & & & & b_{mm} \end{bmatrix}$$

desta forma

$$\text{vecd}(\mathbf{B}) = \begin{bmatrix} b_{11} \\ \vdots \\ b_{ii} \\ \vdots \\ b_{mm} \end{bmatrix} \quad (2.28)$$

Os operadores  $\text{vec}$  e  $\text{vecd}$  são ferramentas úteis na solução de problemas de otimização em redes geodésicas principalmente em problemas sobre a obtenção de pesos em um levantamento geodésico.

Propriedades do operador  $\text{vec}$

$$(1) \text{vec}(\mathbf{A} + \mathbf{B}) = \text{vec}(\mathbf{A}) + \text{vec}(\mathbf{B}); \quad (2.29)$$

$$(2) \text{vec}(\alpha \mathbf{A}) = \alpha \text{vec}(\mathbf{A}), \text{ onde } \alpha \text{ é um número real ;} \quad (2.30)$$

$$(3) \text{tr}(\mathbf{AB}) = \left( \text{vec}(\mathbf{A}^t) \right)^t \text{vec}(\mathbf{B}); \quad (2.31)$$

(4) considerando  $\mathbf{a}$  e  $\mathbf{b}$  como vetores de dimensão  $n \times 1$  e  $m \times 1$  respectivamente, então

$$\text{vec}(\mathbf{ab}^t) = \mathbf{b} \otimes \mathbf{a} \quad (2.32)$$

$$(5) \text{vec}(\mathbf{ABC}) = (\mathbf{C}^t \otimes \mathbf{A}) \text{vec}(\mathbf{B}) \quad (2.33)$$

## 2.8 VALORES E VETORES PRÓPRIOS: ESPECIAL E GENERALIZADO

Um vetor não nulo  $\mathbf{m}$  é dito ser um vetor próprio de uma matriz quadrada  $\mathbf{B}$  se existir um escalar  $\lambda$  tal que,

$$\mathbf{Bm} = \lambda \mathbf{m}. \quad (2.34)$$

Então  $\lambda$  é um valor próprio<sup>1</sup> especial de  $\mathbf{B}$ . Os valores próprios podem ser nulos, no entanto, o vetor próprio não.

A equação de grau  $n$ , expressa por

$$\det(\mathbf{B} - \lambda \mathbf{I}) = 0 \quad (2.35)$$

é denominada equação característica e refere-se a uma matriz  $\mathbf{B}$  de ordem  $n \times n$ .

---

<sup>1</sup> Na literatura podem ser encontradas outras denominações para valor próprio, tais como: autovalor, valor característico, raiz característica, raiz latente, raiz própria, auto-raiz.

Resolvendo a equação característica relativa à  $\lambda$  obtém-se os valores próprios de  $\mathbf{B}$ , que podem ser números reais ou complexos. Se um valor próprio é determinado pode-se substituí-lo na equação (2.34) e esta é resolvida para determinar os vetores próprios correspondentes a cada valor próprio. A expressão polinomial  $\det(\mathbf{B} - \lambda\mathbf{I})$  é também conhecida por polinômio característico de  $\mathbf{B}$ .

Propriedades dos valores e vetores próprios (BRONSON,1993, p.80):

- a) A soma dos valores próprios de uma matriz é igual ao seu traço, que por sua vez é a soma dos elementos da diagonal principal;
- b) Vetores próprios correspondentes a diferentes valores próprios são linearmente independentes;
- c) Uma matriz é dita ser singular se e só se tiver um valor próprio nulo;
- d) Se  $\mathbf{m}$  for um vetor próprio de  $\mathbf{B}$  correspondente ao valor próprio  $\lambda$  e se  $\mathbf{B}$  possuir inversa, então  $\mathbf{m}$  é um vetor próprio de  $\mathbf{B}^{-1}$  correspondente ao valor próprio  $\frac{1}{\lambda}$ ;
- e) Se  $\mathbf{m}$  for um vetor próprio de uma matriz, então  $k \times \mathbf{m}$  também o será para qualquer constante  $k \neq 0$  e  $\mathbf{m}$  ou  $k \times \mathbf{m}$  correspondem ao mesmo valor próprio;
- f) Uma matriz e a sua transposta têm os mesmos valores próprios;
- g) Os valores próprios de uma matriz triangular superior ou inferior são os valores próprios da sua diagonal principal;
- h) O produto de todos os valores próprios de uma matriz é igual ao determinante dessa matriz;
- i) Se  $\mathbf{m}$  for um vetor próprio da matriz  $\mathbf{B}$  correspondente ao valor próprio  $\lambda$ , então  $\mathbf{m}$  é um vetor próprio de  $\mathbf{B} - r\mathbf{I}$  correspondente ao valor próprio  $\lambda - r$  para qualquer escalar  $r$ ;
- j) Os valores próprios de uma matriz simétrica e real são sempre reais;
- l) Os valores próprios de uma matriz definida positiva, ver apêndice, são sempre positivos;

m) Se  $\mathbf{B}$  é simétrica e  $\lambda_i \neq \lambda_j$ , os vetores próprios correspondentes são ortogonais.

Se em vez da matriz identidade em (2.35) tiver uma matriz  $\mathbf{C}$ , haverá como ampliação do problema de valor próprio especial o problema de valor próprio geral definido por:

$$(\mathbf{B} - \lambda\mathbf{C})\mathbf{m} = \mathbf{0} \quad (2.36)$$

em que  $\mathbf{B}$  e  $\mathbf{C}$  possuem as mesmas dimensões,  $n \times n$ . Em particular, no caso em que  $\mathbf{C}$  for não-singular será possível reconduzir (2.36) a um problema de valor próprio especial. Com efeito, pré-multiplicando ambos os membros da equação (2.36) por  $\mathbf{C}^{-1}$

$$(\mathbf{C}^{-1}\mathbf{B} - \lambda\mathbf{C}^{-1}\mathbf{C})\mathbf{m} = \mathbf{0} \quad (2.37)$$

e fazendo  $\mathbf{C}^{-1}\mathbf{B} = \mathbf{D}$  e ainda como  $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I}$  tem-se

$$(\mathbf{D} - \lambda\mathbf{I})\mathbf{m} = \mathbf{0} \quad (2.38)$$

a qual coincide com a definição de um problema de valor próprio especial.

### 3 MÉTODOS ANALÍTICOS BASEADOS EM MATRIZ CRITÉRIO DE COVARIÂNCIAS PARA A OTIMIZAÇÃO DOS PESOS DAS OBSERVAÇÕES GEODÉSICAS

Esta seção trata da determinação dos pesos das observações geodésicas sob o ponto de vista dos métodos analíticos especificamente quando se faz uso de modelos idealizados em um sentido ótimo para a matriz de covariâncias dos parâmetros estimados.

Especificamente em um projeto de segunda ordem, as equações para sua solução iniciam da expressão geral para a matriz de covariâncias dos parâmetros estimados para os pontos da rede:

$$\mathbf{Q}_x = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \quad (3.1)$$

Usando a pseudo-inversa ou inversa generalizada de Moore-Penrose em ambos os membros de (3.1) obtém-se

$$\mathbf{Q}_x^+ = \mathbf{A}^t \mathbf{P} \mathbf{A} = \mathbf{N} \quad (3.2)$$

onde  $\mathbf{A}$  é a matriz de planejamento ou matriz de configuração da rede,  $\mathbf{P}$  é a matriz dos pesos das observações e  $\mathbf{N}$  é a matriz dos coeficientes das equações normais. No caso de um planejamento de redes geodésicas, com base em matriz critério, o símbolo  $\mathbf{N}$  pode ser substituído pela letra  $\mathbf{Q}$  para representar a inversa de uma matriz critério de covariâncias.

Como as informações a respeito da precisão de uma rede geodésica estão contidas na matriz de covariâncias dos parâmetros estimados, todas as modificações na geometria da rede e na precisão das observações agem diretamente sobre (3.1), desta forma os critérios de análise recaem sobre esta matriz.

As matrizes  $\mathbf{A}$  e  $\mathbf{Q}$  são conhecidas a priori.  $\mathbf{A}$  é a matriz de planejamento da rede e traz as informações sobre o tipo de observação que será realizada e  $\mathbf{N}$  é a inversa da matriz critério de covariância  $\mathbf{Q}_x$  que estabelece a precisão idealizada para as coordenadas estimadas da rede. Com isso pode-se formular a solução para a otimização baseando-se na equação (3.2) a qual permite a melhor aproximação possível da matriz critério.

Neste sentido a matriz dos pesos pode ser considerada como:

a) Problema semi-defnido positivo:

considera a matriz dos pesos “cheia” , ou seja, os elementos fora da diagonal principal não são todos nulos.

b) Problema diagonal:

considera a matriz dos pesos uma matriz diagonal, ou seja, os elementos fora da diagonal principal são todos nulos.

### 3.1 PROBLEMA SEMI-DEFINIDO POSITIVO

O problema definido positivo considera a matriz dos pesos  $\mathbf{P}$  como uma matriz completa, isto implica que as observações são correlacionadas. Este fato não garante que esta matriz possa ser realizada na prática visto que o correlacionamento entre as observações não é um fato perfeitamente compreendido nos problemas de ajustamento em redes geodésicas.

Aqui são apresentadas duas possibilidades para a solução do caso da matriz dos pesos completa.

#### 3.1.1 Solução com o Produto de Kronecker

Esta solução é obtida com o uso da propriedade (2.33) do operador vec na expressão (3.2), originando o sistema de equações,

$$\left(\mathbf{A}^t \otimes \mathbf{A}^t\right) \mathbf{p} = \mathbf{q} \quad (3.3)$$

cuja solução é representado por

$$\mathbf{p} = \left(\mathbf{A}^t \otimes \mathbf{A}^t\right)^+ \mathbf{q} \quad (3.4)$$

com  $\mathbf{p} = \text{vec}(\mathbf{P})$  vetor dos elementos da matriz  $\mathbf{P}$ , dispostos coluna sobre coluna ordenadamente e  $\mathbf{q}$  vetor dos elementos da matriz  $\mathbf{Q}$  (lembrando que  $\mathbf{Q} \equiv \mathbf{N}$ ), dispostos coluna sobre coluna ordenadamente.

### 3.1.2 Solução para a matriz completa dos pesos com auxílio direto da Pseudo-Inversa

Esta segunda possibilidade obtém uma matriz dos pesos completa com o uso da pseudo-inversa na equação (3.2). Tal solução tem a forma:

$$\mathbf{P} = \left( \mathbf{A}^+ \right)^t \mathbf{Q}_x^t \mathbf{A}^+ \quad (3.5)$$

a qual produz uma matriz dos pesos completa.

## 3.2 PROBLEMA DIAGONAL

Para o problema diagonal onde se considera as observações como não correlacionadas, a matriz  $\mathbf{P}$  terá estrutura diagonal. Para este caso há basicamente três modelos. Dois desses modelos fazem uma aproximação direta da matriz critério de covariâncias e o outro aproxima a inversa desta matriz, representada por  $\mathbf{Q}$ . Os três modelos são descritos a seguir de forma resumida e para maiores informações deve-se consultar, por exemplo, SCHMITT (1985).

### 3.2.1 Aproximação Direta da Matriz Critério de Covariância

Partindo da expressão (3.2)

$$\mathbf{A}^t \mathbf{P} \mathbf{A} = \mathbf{Q}_x^+ \quad (3.6)$$

pré e pós multiplicando a equação (3.6) pela matriz critério  $\mathbf{Q}_x$  obtém-se

$$\mathbf{Q}_x \mathbf{A}^t \mathbf{P} \mathbf{A} \mathbf{Q}_x = \mathbf{Q}_x \mathbf{Q}_x^+ \mathbf{Q}_x = \mathbf{Q}_x \quad (3.7)$$

e a equação (3.7), fazendo  $\mathbf{K} = \mathbf{Q}_x \mathbf{A}^t$ , escreve-se

$$\mathbf{K} \mathbf{P} \mathbf{K}^t = \mathbf{Q}_x \quad (3.8)$$

O problema pode ser solucionado para  $\mathbf{P}$  utilizando o produto Khatri-Rao e a propriedade (2.33) aplicado ao sistema de equações (3.8) (RAO; MITRA, 1971, p. 11)

$$(\mathbf{K} \odot \mathbf{K}) \mathbf{p} = \mathbf{q}_x \quad (3.9)$$

na equação (3.9) tem-se que  $\odot$  representa o operador do produto de Katri-Rao,  $\mathbf{p} = \text{vecd}(\mathbf{P})$  e  $\mathbf{q}_x = \text{vec}(\mathbf{Q}_x)$ .

A solução geral de (3.9), com o uso da pseudo-inversa, é escrita como,

$$\mathbf{p} = (\mathbf{K} \odot \mathbf{K})^+ \mathbf{q}_x. \quad (3.10)$$

considerando  $n$  como número de observações e  $u$  número de parâmetro a dimensão de  $(\mathbf{K} \odot \mathbf{K})^+$  e  $\mathbf{q}_x$  são, respectivamente,  $n \times u^2$  e  $u^2 \times 1$ . Desta forma a dimensão de  $\mathbf{p}$  é  $n \times 1$  confirmando que está solução considera a matriz de covariâncias das observações como não correlacionada fato que é usual nas Ciências Geodésicas.

### 3.2.2 Aproximação Iterativa da Matriz Critério de Covariância

A partir da lei de propagação das covariâncias, aplicada à solução de um ajustamento de observações pelo método dos mínimos quadrados na forma paramétrica (GEMAEL, 1994, p.119)

$$\mathbf{x} = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \mathbf{A}^t \mathbf{P} \ell \quad (3.11)$$

que conduz a expressão da matriz de covariâncias dos parâmetros estimados

$$\mathbf{Q}_x = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \mathbf{A}^t \mathbf{P} \Sigma_{\ell_b} \mathbf{P} \mathbf{A} (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \quad (3.12)$$

onde  $\Sigma_{\ell_b}$  é a matriz de covariâncias das observações.

Na (3.12) fazendo  $\mathbf{H} = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \mathbf{A}^t \mathbf{P}$  esta é reescrita como

$$\mathbf{H} \Sigma_{\ell_b} \mathbf{H}^t = \mathbf{Q}_x \quad (3.13)$$

com o uso do produto de Khatri-Rao a equação (3.13) torna-se,

$$(\mathbf{H} \odot \mathbf{H}) \mathbf{s} = \mathbf{q}_x \quad (3.14)$$

onde  $\mathbf{s} = \text{vecd}(\Sigma_{\ell_b})$  e  $\mathbf{q}_x = \text{vec}(\mathbf{Q}_x)$ .

Nestas condições a equação (3.14) deve ser resolvida iterativamente, isto por que a matriz  $\mathbf{H}$  contém o valor atual da matriz dos pesos  $\mathbf{P}$  e esta é atualizada pela

nova matriz  $\mathbf{P} = \boldsymbol{\Sigma}_{\ell_b}^+$ . Considerando o processo iterativo que conduz à solução do problema tem-se o seguinte algoritmo

Para  $k = 0$  escolhe-se um valor inicial para a matriz dos pesos e um critério de parada, por exemplo  $\mathbf{P}^0 = \mathbf{I}$ ,  $\varepsilon = 10^{-8}$ , respectivamente, e faz-se para  $k = 0, 1, 2, 3, \dots$

1. forma-se a matriz  $\mathbf{H}^k = (\mathbf{A}^t \mathbf{P}^k \mathbf{A})^+ \mathbf{A}^t \mathbf{P}^k$

2. obtém-se

$$\mathbf{s}^{k+1} = (\mathbf{H}^k \odot \mathbf{H}^k)^+ \mathbf{q}_x$$

3. constrói-se a matriz dos pesos  $\mathbf{P}^{k+1}$  fazendo  $\left(\boldsymbol{\Sigma}_{\ell_b}^{k+1}\right)^+ = \text{diag}(\mathbf{s}^{k+1})$  e define-se

$$\mathbf{P}^{k+1} = \left(\boldsymbol{\Sigma}_{\ell_b}^{k+1}\right)^+$$

4. se  $\|\mathbf{s}^{k+1} - \mathbf{s}^k\| \geq \varepsilon$  faz-se

$k = k + 1$  e volta ao passo 1, caso contrário o processo é interrompido.

### 3.2.3 Aproximação Direta da Inversa da Matriz Critério de Covariância.

Um terceiro modelo que é utilizado para obtenção dos pesos, baseia-se na matriz inversa da matriz critério de covariâncias representada por  $\mathbf{Q}$ , ou seja,

$$\mathbf{Q} = \mathbf{Q}_x^+$$

Desta forma pode-se utilizar o produto de Khatri-Rao diretamente sobre a equação (3.6) obtendo-se,

$$(\mathbf{A}^t \odot \mathbf{A}^t) \mathbf{p} = \mathbf{q} \quad (3.15)$$

na equação (3.15)  $\mathbf{p} = \text{vecd}(\mathbf{P})$  e  $\mathbf{q} = \text{vec}\left(\mathbf{Q}_x^+\right)$  cuja solução geral é representada por,

$$\mathbf{p} = (\mathbf{A}^t \odot \mathbf{A}^t)^+ \mathbf{q} \quad (3.16)$$

Em qualquer caso pode ser utilizado como medida para verificar a qualidade da aproximação, obtida pela aplicação de um dos métodos anteriormente citados, a

expressão dada pela diferença entre a matriz de covariâncias e a matriz critério de covariância utilizada no método, isto é,

$$\mathbf{D} = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ - \mathbf{Q}_x \quad (3.17)$$

na qual  $\mathbf{P} = \text{diag}(\mathbf{p})$  é o resultado de uma das soluções obtidas pelo problema semi-definido positivo ou problema diagonal

Chamando  $\mathbf{D}$  de matriz residual e fazendo  $\mathbf{d} = \text{vec}(\mathbf{D})$  pode-se utilizar como medida global para a qualidade da aproximação a relação,

$$\mathbf{d}^t \mathbf{d} \quad (3.18)$$

que é a soma dos quadrados dos resíduos referidos à matriz  $\mathbf{D}$ .

### 3.3 EXEMPLIFICAÇÃO PARA OS MÉTODOS BASEADOS EM MATRIZ CRITÉRIO

Com a finalidade de ilustrar a aplicação dos métodos baseados em matriz critério e obter as grandezas necessárias para sua aplicação apresenta-se um exemplo simples aplicado a uma rede de nivelamento.

Na figura 3.1 a altura do ponto h é zero e por simplicidade considerada isenta de erro.

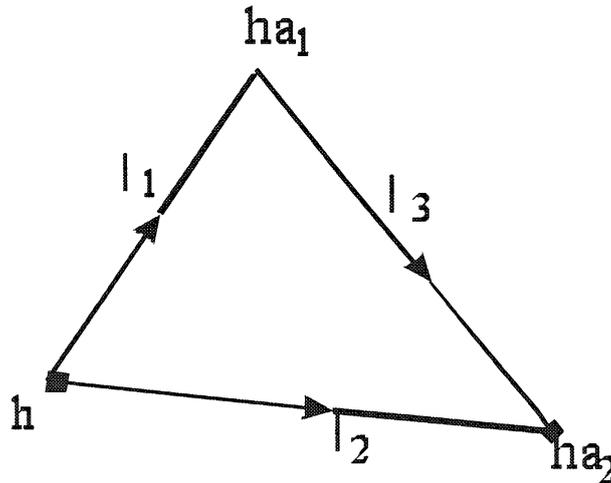
Com as seguintes equações de observação:

$$\ell_1 = ha_1$$

$$\ell_2 = ha_2$$

$$\ell_3 = ha_2 - ha_1$$

FIGURA 3.1 - REDE DE NIVELAMENTO



A matriz planejamento é:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

O número de observações planejadas é 3 e o número de parâmetros (desníveis) à serem estimados é dois, com isso a matriz critério de covariância terá dimensão  $2 \times 2$ . Deseja-se que a precisão de cada parâmetro seja a unidade e que a correlação entre os parâmetros seja 0,5 com isso a matriz critério é escrita como,

$$\mathbf{Q}_x = \begin{bmatrix} 1 & 0,5 \\ 0,5 & 1 \end{bmatrix}.$$

Para uso nos modelos de otimização dos pesos se faz necessária a forma vetorial da matriz  $\mathbf{Q}_x$  desta forma utilizando o operador  $\text{vec}$  tem-se,

$$\mathbf{q}_x = \text{vec}(\mathbf{Q}_x) = \begin{bmatrix} 1 \\ 0,5 \\ 0,5 \\ 1 \end{bmatrix}.$$

Em alguns casos, necessita-se da inversa da matriz critério representada por  $\mathbf{Q}$  cujo valor é  $\mathbf{Q} = \begin{bmatrix} 1,333333 & -0,666666 \\ -0,666666 & 1,333333 \end{bmatrix}$ .

A forma vetorial da matriz  $\mathbf{Q}$ , com o uso do operador  $\text{vec}$  é dada por:

$$\mathbf{q} = \text{vec}(\mathbf{Q}) = \begin{bmatrix} 1,333333 \\ -0,666666 \\ -0,666666 \\ 1,333333 \end{bmatrix}.$$

Desta forma enuncia-se o seguinte problema: qual peso deve ser atribuído à cada observação para que a rede apresente uma matriz de covariâncias como a estabelecida por  $\mathbf{Q}_x$ ?

A solução para o problema semi-definido positivo e problema diagonal e ainda suas variações estão colocadas à seguir.

### 1) Obtenção de uma matriz dos pesos completa

Para utilização da equação (3.3) se faz necessário da forma vetorial da matriz inversa da matriz critério de covariâncias e do produto de Kronecker ( $\mathbf{A}^t \otimes \mathbf{A}^t$ ). O sistema linear escrito explicitamente torna-se,

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{bmatrix} \times \text{vec}(\mathbf{P}) = \begin{bmatrix} 1,333333 \\ -0,666666 \\ -0,666666 \\ 1,333333 \end{bmatrix}$$

Para esta solução fez-se uso também do modelo dado na equação (3.5) e a solução encontrada em ambos os casos foi,

$$\mathbf{P} = \begin{bmatrix} 0,4444444 & 0,2222222 & -0,2222222 \\ 0,2222222 & 0,4444444 & 0,2222222 \\ -0,2222222 & 0,2222222 & 0,4444444 \end{bmatrix}.$$

Como exposto anteriormente, este tipo de solução tem sua aplicação prática reduzida devido ao pouco conhecimento que se tem de uma matriz de covariâncias das observações correlacionadas (matriz dos pesos completa).

### 2) Obtenção da matriz dos pesos diagonal.

#### a) Modelo pela aproximação direta da matriz critério de covariâncias

Para este caso faz-se uso da equação (3.9) o que necessita da forma vetorial da matriz critério de covariâncias  $\mathbf{q}_x = \text{vec}(\mathbf{Q}_x)$  (e não de sua inversa). Do produto,

$$\mathbf{K} = \mathbf{Q}_x \mathbf{A}^t = \begin{bmatrix} 1,0 & 0,5 & -0,5 \\ 0,5 & 1 & 0,5 \end{bmatrix}$$

e do produto de Khatri-Rao ( $\mathbf{K} \odot \mathbf{K}$ ). O sistema linear escrito explicitamente torna-se

$$\begin{bmatrix} 1,00 & 0,25 & 0,25 \\ 0,50 & 0,50 & -0,25 \\ 0,50 & 0,50 & -0,25 \\ 0,25 & 1,00 & 0,25 \end{bmatrix} \times \begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \\ p_{44} \end{bmatrix} = \begin{bmatrix} 1 \\ 0,5 \\ 0,5 \\ 1 \end{bmatrix}.$$

O uso da solução (3.10) com o auxílio da inversa generalizada de Moore-Penrose conduziu à seguinte solução apresentada em forma de matriz,

$$\mathbf{P} = \begin{bmatrix} 0,6666667 & 0 & 0 \\ 0 & 0,6666667 & 0 \\ 0 & 0 & 0,6666667 \end{bmatrix}.$$

Esta solução tem uma melhor aplicação prática em relação à matriz completa, fica condicionada ao instrumental que pode conduzir à variâncias compatíveis com a matriz peso alcançada. A matriz de covariâncias das observações, compatível com a matriz  $\mathbf{P}$ , é dada por,

$$\Sigma_{\ell_b} = \mathbf{P}^{-1} = \begin{bmatrix} 1,5 & 0 & 0 \\ 0 & 1,5 & 0 \\ 0 & 0 & 1,5 \end{bmatrix},$$

ou seja, cada linha do nivelamento deve ser feita com uma precisão melhor ou igual à 1,23 aproximadamente para que se cumpra a postulação de precisão dada pela matriz critério de covariâncias.

b) Modelo pela aproximação iterativa da matriz critério de covariâncias.

Este método usa diretamente a forma vetorial da matriz critério de covariâncias  $\mathbf{q}_x = \text{vec}(\mathbf{Q}_x)$  como no caso anterior. Necessita-se ainda do produto

$\mathbf{H} = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+ \mathbf{A}^t \mathbf{P}$  que para ser escrito explicitamente é arbitrado um valor inicial para a matriz  $\mathbf{P}$ , visto que este método se resolve por um processo iterativo. A matriz  $\mathbf{P}$

utilizada para iniciar o processo iterativo é a matriz identidade de dimensão  $3 \times 3$ , ou seja,  $\mathbf{P}^0 = \mathbf{I}$ . Com isso,

$$\mathbf{H}^0 = \begin{bmatrix} 0,666666 & 0,333333 & -0,333333 \\ 0,333333 & 0,666666 & 0,333333 \end{bmatrix}$$

e o sistema linear (3.14) escrito explicitamente é

$$\begin{bmatrix} 0,444444 & 0,111111 & 0,111111 \\ 0,222222 & 0,222222 & -0,111111 \\ 0,222222 & 0,222222 & -0,111111 \\ 0,111111 & 0,444444 & 0,111111 \end{bmatrix} \times \text{vecd}(\boldsymbol{\Sigma}_{\ell_b}) = \begin{bmatrix} 1 \\ 0,5 \\ 0,5 \\ 1 \end{bmatrix}$$

cuja solução com o uso da pseudo-inversa é:

$$\mathbf{s}^1 = \text{vecd}(\boldsymbol{\Sigma}_{\ell_b}) = \begin{bmatrix} 1,5 \\ 1,5 \\ 1,5 \end{bmatrix}.$$

O processo iterativo necessitou somente de uma iteração para alcançar a solução, isto se deve ao fato do valor arbitrado para a matriz  $\mathbf{P}$ , já ser um valor suficientemente bom.

Uma característica deste método é que ele obtém em cada passo diretamente, não a matriz dos pesos e sim a matriz de covariâncias das observações, porém para reiniciar em cada passo, o processo iterativo, é necessário a matriz dos pesos, obtida por  $\mathbf{P} = \boldsymbol{\Sigma}_{\ell_b}^{-1}$ , condição necessária para calcular a matriz  $\mathbf{H}$ .

Desta forma a matriz dos pesos resultante é

$$\mathbf{P} = \boldsymbol{\Sigma}_{\ell_b}^{-1} = \begin{bmatrix} 0,666666 & 0 & 0 \\ 0 & 0,666666 & 0 \\ 0 & 0 & 0,666666 \end{bmatrix}$$

a qual conduz a precisão pré-definida no planejamento.

c) Modelo pela aproximação direta da inversa da matriz critério de covariâncias

Utilizando o produto Khatri-Rao (2.24) na equação (3.6), o sistema (3.15) considerando a forma vetorial da inversa da matriz critério de covariâncias  $\mathbf{Q}$  dada por  $\mathbf{q} = \text{vec}(\mathbf{Q})$ . O sistema linear é

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \\ p_{44} \end{bmatrix} = \begin{bmatrix} 1,333333 \\ -0,666666 \\ -0,666666 \\ 1,333333 \end{bmatrix}.$$

Desta forma a solução do sistema de equações lineares, estabelecido na (3.16) apresenta a seguinte matriz dos pesos como solução:

$$\mathbf{P} = \begin{bmatrix} 0,666666 & & \\ & 0,666666 & \\ & & 0,666666 \end{bmatrix}$$

a qual como nos casos anteriores leva à precisão pré-definida pela matriz critério de covariâncias.

#### 4 A OTIMIZAÇÃO DOS PESOS DAS OBSERVAÇÕES GEODÉSICAS BASEADA EM VALORES PRÓPRIOS

Uma outra possibilidade de se obter pesos otimizados, é o uso dos valores próprios associados à matriz de covariâncias dos parâmetros ajustados, em termos das Componentes Principais, os quais possuem informações sobre a qualidade da rede. O objetivo é estabelecer valores próprios que conduzam a matriz de covariâncias à apresentação de uma estrutura espectral<sup>2</sup> ideal e de forma que possua os valores próprios requeridos.

O valor próprio máximo da matriz de covariâncias dos parâmetros estimados é de interesse particular neste processo de otimização, pois este permite obter informação sobre uma precisão limite para quaisquer grandezas estimadas a partir da rede (NINKOV; SCHMITT, 1983, p. 217). Com efeito para uma função  $y$  das coordenadas, após linearização pela expansão de Taylor vale,

$$\mathbf{y} = \mathbf{f}^t \mathbf{x} \quad (4.1)$$

e para a sua variância após aplicação da lei de propagação de covariâncias tem-se

$$\sigma_y^2 = \sigma_0^2 \mathbf{f}^t \mathbf{Q}_x \mathbf{f}. \quad (4.2)$$

Com o auxílio do Quociente Rayleigh (WANG; CHOW, 1994, p. 38) obtém-se uma estimativa para  $\sigma_y^2$  pela expressão

$$\mathbf{f}^t \mathbf{f} \lambda_{\min} \leq \sigma_y^2 \leq \mathbf{f}^t \mathbf{f} \lambda_{\max} \quad (4.3)$$

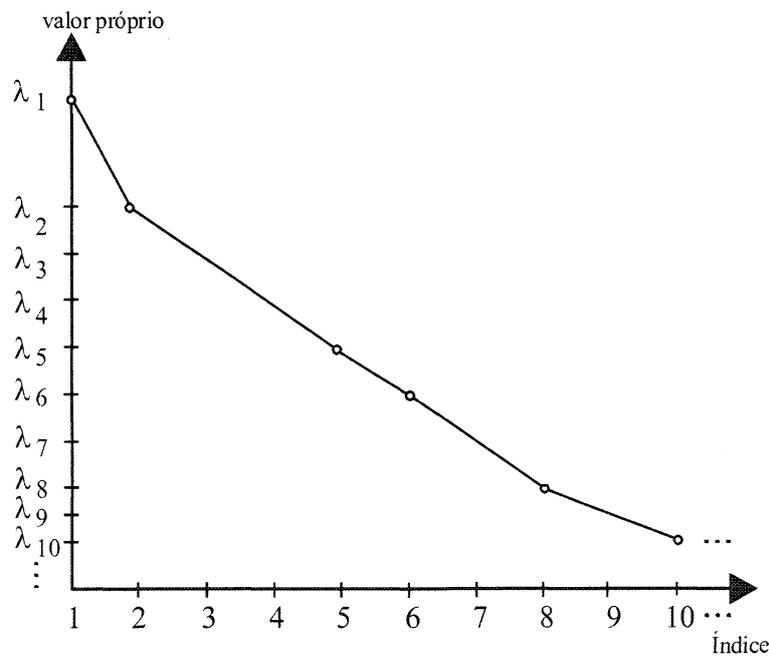
com  $\lambda_{\max}$  e  $\lambda_{\min}$  obtidos de (3.1).

Da (4.3) decorre a exigência que o valor próprio máximo da matriz de covariâncias dos parâmetros estimados deve ser mínimo. Quanto maior for um valor próprio, comparativamente em relação aos outros elementos do espectro, mais desfavorável e não homogêneo será o comportamento da precisão como mostra a figura 4.1. O ideal é que o espectro da matriz de covariâncias seja o mais homogêneo possível.

---

<sup>2</sup> Espectro: conjunto dos valores próprios de uma matriz

FIGURA 4.1 – EXEMPLO DA REPRESENTAÇÃO GRÁFICA DO ESPECTRO DE UMA MATRIZ



Com o problema de valor próprio inverso aplicado à otimização dos pesos, quer-se garantir que qualquer desvio padrão estimado a partir da rede não cruze um valor limite pretendido, com isso o valor próprio máximo da matriz  $\mathbf{Q}_x$  pode ser fixado e tido como um valor limite  $\lambda_{\max}^*$  (valor próprio máximo pretendido), o qual se torna o objetivo no processo de otimização dos pesos modelado por um problema de valor próprio máximo.

Com a definição de um valor limite superior para a precisão dos parâmetros tem-se o valor próprio máximo que é pretendido para a matriz de covariâncias.

A tomada de decisão sobre os valores próprios que devem ser estabelecidos em um processo de otimização é uma tarefa importante. Desta forma algumas ferramentas podem ser úteis para auxiliar na decisão correta, como os critérios de otimalidade para redes geodésicas e o teste para igualdade de um conjunto dos valores próprios de uma matriz de covariâncias.

#### 4.1 CRITÉRIOS DE OTIMALIDADE PARA REDES GEODÉSICAS

Em geral, os critérios de otimalidade para precisão são descritos pelas medidas de precisão usadas para descrever a qualidade da rede. Na fase de planejamento da rede geodésica é possível avaliar o tamanho e a forma do elipsóide de confiança, visto que estas grandezas são dependentes apenas dos valores próprios e vetores próprios da matriz  $\mathbf{Q}_x$  (PELZER, 1980, p. 57) .

No caso onde os requerimentos de precisão não são facilmente definidos, alguns conceitos de rede ideal são propostos.

Alguns dos mais importantes critérios de otimalidade e exigências para a rede geodésica (DUPRAZ; NIEMEIER, 1981, p. 387-389; WELSH et al., 2000, p.133-134)

são (onde nas equações (4.4) a (4.8) o símbolo  $!$  indica: deve ser) :

$$\prod_{i=1}^p \lambda_i = \det(\mathbf{Q}_x) = \lambda_1 \times \lambda_2 \times \dots \times \lambda_p \stackrel{!}{=} \text{mín.} \quad (4.4)$$

é a medida denominada variância generalizada que deve ser mínima;

$$\sum_{i=1}^p \lambda_i = \text{tr}(\mathbf{Q}_x) = \lambda_1 + \lambda_2 + \dots + \lambda_p \stackrel{!}{=} \text{mín.} \quad (4.5)$$

é a medida denominada variância total e deve ser mínima;

$$\lambda_{\max} \stackrel{!}{=} \text{mín.} \quad (4.6)$$

significa que o quadrado do semi-eixo maior da elipse de erros deve ser mínimo, indicando que a precisão de uma rede será tanto mais alta quanto menor for o valor próprio máximo da matriz de covariância do vetor dos parâmetros estimados  $\hat{x}$ ;

$$\frac{\lambda_{\max}}{\lambda_{\min}} \stackrel{!}{=} 1 \quad (4.7)$$

é conhecida como a condição de isotropia, ou seja, a medida da precisão do ponto é a mesma em todas as direções;

$$\lambda_{\max} - \lambda_{\min} \stackrel{!}{=} \text{mín.} \quad (4.8)$$

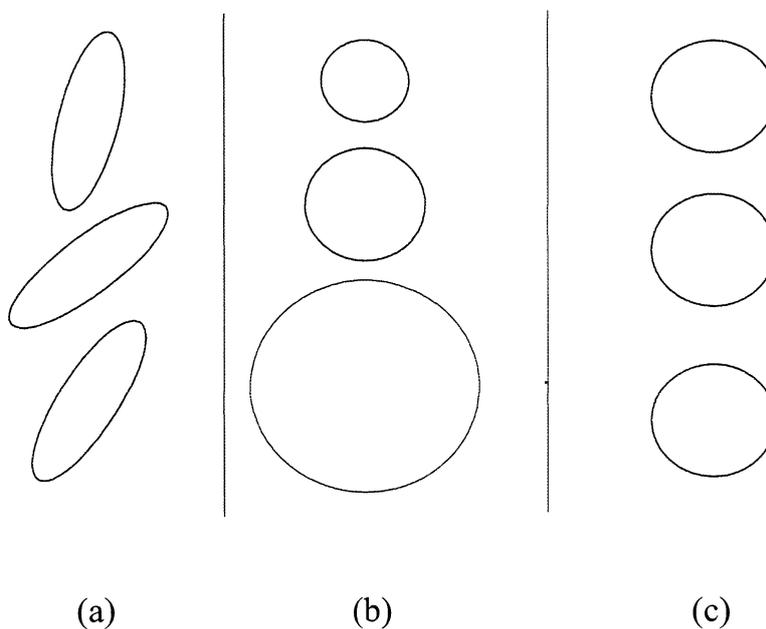
é a condição de homogeneidade, ou seja, as elipses se aproximam de uma circunferência e

$$p_1 = \mathbf{m} \sqrt{\lambda_{\max.}} \quad (4.9)$$

fornece a direção e o comprimento do semi-eixo maior do elipsóide de confiança em termos da primeira componente principal.

Uma rede que é somente homogênea (figura 4.2 a), as elipses (ou elipsóides) de erro locais são os mesmos em todos os pontos. Uma rede que é apenas isotrópica (figura 4.2 b), as elipses (ou elipsóides) variam de ponto para ponto, embora sejam todas reduzidas a círculos (redes bidimensionais) ou esferas (redes tridimensionais). Assumindo-se uma rede geodésica bidimensional como sendo homogênea e isotrópica (figura 4.2 c), então as elipses de erro locais reduzem-se a círculos de mesmo raio.

FIGURA 4.2 – REDE HOMOGÊNEA E ISOTRÓPICA



#### 4.2 TESTE PARA A IGUALDADE DE VALORES PRÓPRIOS

A fim de verificar, se sob um nível de significância  $\alpha$ , existe um conjunto de  $p$  valores próprios iguais entre si utiliza-se o teste da igualdade de valores próprios. Este procedimento conduz a verificação se os elementos de um conjunto de valores próprios são significativamente distintos ou iguais entre si. Considerando o conjunto

$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$ . O teste pode ser usado para qualquer subconjunto consecutivo de valores próprios. Se um subconjunto de  $b$  valores próprios é dado,  $\lambda_{k+1} \leq \lambda_{k+2} \leq \dots \leq \lambda_{k+b}$ , a hipótese nula é (JACKSON, 1991, p.86-87):

$$\mathbf{H}_0 : \lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_{k+b}$$

Para esta hipótese, a estatística a ser calculada é,

$$v \left[ - \sum_{j=k+1}^{k+b} \ln(\lambda_j) + b \ln \left( \sum_{j=k+1}^{k+b} \frac{\lambda_j}{b} \right) \right] \sim \chi^2, \quad (4.5)$$

onde  $v$  denota o número de graus de liberdade associado com a matriz de covariâncias e a distribuição  $\chi^2$  tem  $(b-1)(b+2)/2$  graus de liberdade. Fixado o nível de significância  $\alpha$ , se a estatística calculada for maior que a estatística  $\chi_{(b-1)(b+2)/2}^2$ , rejeita-se a hipótese nula.

Para o caso em que  $p=2$ , caso bivariado, o teste da igualdade dos valores próprios sob a hipótese nula é:

$$\mathbf{H}_0 : \lambda_1 = \lambda_2$$

e para esta hipótese, a estatística a ser calculada é,

$$F^* = \frac{(n-2)(\lambda_1 - \lambda_2)^2}{8\lambda_1\lambda_2} \quad (4.6)$$

Na relação (4.6) a estatística  $F^*$  a ser testada tem distribuição F central com número de graus de liberdade no numerador igual a 2 e o número de graus de liberdade no denominador  $n-2$ , ou seja,  $F^* \sim F_{2, n-2}$  e  $n$  é o número de observações.

Este teste permite formular hipóteses sobre uma rede geodésica, como exemplo, pode ser feita a inferência se uma rede é homogênea, ou não, sob um nível de significância. Para o caso de planejamento de redes o teste pode ser utilizado antes mesmo de qualquer medição em campo, ou seja, pode ser aplicado ao conjunto de valores próprios pré-estabelecidos que serão utilizados no processo de otimização.

Visto que  $\mathbf{N}$  e  $\mathbf{Q}_x$  apresentam valores próprios recíprocos entre si, vale com respeito ao valor próprio  $\lambda_i$  de  $\mathbf{N}$  e  $\mu_i$  de  $\mathbf{Q}_x$  a relação

$$\lambda_i = \frac{1}{\mu_i} \quad \text{com } \mu_i \neq 0 \text{ e } i = 1, 2, \dots, u \quad (4.7)$$

onde  $u$  é o número de parâmetros.

Com isso há a possibilidade de ser formulada a otimização relativa à matriz das equações normais  $\mathbf{N}$ . Com base em (4.7) verifica-se que um aumento no valor próprio  $\lambda_i$  de  $\mathbf{N}$  corresponde a um decréscimo no valor próprio  $\mu_i$  de  $\mathbf{Q}_x$ , ou seja, para  $\mu_i$  tendendo para seu valor máximo tem-se  $\lambda_i$  tendendo para seu valor mínimo. Todo o processo aplicado para a obtenção dos valores próprios pretendidos é efetuado sobre a matriz dos coeficientes das equações normais. Então, definido o valor limite para a precisão dos parâmetros  $\mu_{\max}$  obtém-se os valores próprios para a matriz  $\mathbf{N}$  através da relação (4.7).

## 5 O PROBLEMA DE VALOR PRÓPRIO INVERSO

Um problema de valor próprio inverso diz respeito à construção de uma matriz com dados espectrais pré-estabelecidos. Os dados espectrais podem consistir de informações parciais ou completas dos valores próprios ou dos vetores próprios. O objetivo principal de um problema de valor próprio inverso é obter uma matriz que preserve as propriedades espectrais pré-definidas.

No problema de valor próprio determinam-se os valores próprios para uma matriz quadrada. Em se tratando de valor próprio inverso o problema consiste em determinar uma matriz que possua valores próprios pretendidos, ou seja, se o problema é direto os valores próprios serão incógnitas e no problema inverso a matriz será a incógnita. De uma forma similar também se pode conceituar um problema de vetor próprio inverso.

Para conceituar um problema de valor próprio inverso (PVPI), considera-se a seguinte situação:

Seja a matriz  $\mathbf{A}(\mathbf{c})$ , de ordem  $n$ , definida como,

$$\mathbf{A}(\mathbf{c}) = \mathbf{A}_0 + \sum_{k=1}^n c_k \mathbf{A}_k \quad (5.1)$$

onde  $\mathbf{A}_0$  e  $\mathbf{A}_k$  são matrizes reais e simétricas de dimensão  $n \times n$  e

$\mathbf{c} = [c_1 \ c_2 \ \dots \ c_j \ \dots \ c_n]^t$  um vetor que pertence ao espaço  $n$ -dimensional dos

números reais  $\mathbb{R}^n$ . Denotando o conjunto dos  $n$  valores próprios de  $\mathbf{A}(\mathbf{c})$  por

$\lambda(\mathbf{c}) = \{\lambda_1(\mathbf{c}), \lambda_2(\mathbf{c}), \dots, \lambda_i(\mathbf{c}), \dots, \lambda_n(\mathbf{c})\}$ , com os elementos ordenados de forma

crescente,  $\lambda_1(\mathbf{c}) \leq \lambda_2(\mathbf{c}) \leq \dots \leq \lambda_i(\mathbf{c}) \leq \dots \leq \lambda_n(\mathbf{c})$  encontrar o vetor  $\mathbf{c}$  que leva a matriz

(5.1) à apresentar os valores próprios pré-estabelecidos

$\lambda^* = [\lambda_1^* \ \lambda_2^* \ \dots \ \lambda_i^* \ \dots \ \lambda_n^*]^t$  cujos elementos também estão em ordem crescente.

Desta forma espera-se que para algum vetor  $\mathbf{c} = \mathbf{c}^*$ , se encontre a igualdade

$$\lambda(\mathbf{c}^*) - \lambda^* = \mathbf{0}.$$

Estas considerações definem um problema de valor próprio inverso.

Segundo DOWNING e HOUSEHOLDER (1956, p. 203), os problemas de valor próprio inverso mais comumente encontrados são idealizados com as seguintes proposições:

- (i) encontrar uma matriz  $\mathbf{D}$  tal que o espectro de  $\mathbf{A} + \mathbf{D}$  seja um conjunto pré-fixado de valores próprios  $\lambda = \{\lambda_1, \dots, \lambda_n\} \in \mathbb{C}$ , onde  $\mathbb{C}$  representa o conjunto dos números complexos;
- (ii) encontrar uma matriz  $\mathbf{D}$  tal que o espectro de  $\mathbf{AD}$  seja um conjunto pré-estabelecido de valores próprios  $\lambda = \{\lambda_1, \dots, \lambda_n\} \in \mathbb{C}$ .

O primeiro caso, é chamado de problema de valor próprio inverso aditivo e o segundo, de problema de valor próprio inverso multiplicativo. Frequentemente em aplicações práticas  $\mathbf{A}$ ,  $\mathbf{D}$  e  $\lambda$ , como definidos acima, são grandezas reais.

Na área das Ciências Geodésicas, o interesse pelo problema de valor próprio inverso é motivado pelo planejamento de Redes Geodésicas, mais especificamente na obtenção dos pesos para as observações geodésicas.

## 5.1 O PROBLEMA DE VALOR PRÓPRIO INVERSO ADITIVO

Definição:

Dados, matrizes  $n \times n$  reais e simétricas  $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_k, \dots, \mathbf{A}_n$  e números reais, encontrar o vetor  $\mathbf{c} \in \mathbb{R}^n$  de tal forma que os valores próprios  $\lambda_1(\mathbf{c}) \leq \lambda_2(\mathbf{c}) \leq \dots \leq \lambda_i(\mathbf{c}) \leq \dots \leq \lambda_n(\mathbf{c})$  da equação (5.1) satisfaçam a igualdade  $\lambda_i(\mathbf{c}) = \lambda_i^*$  para  $i = 1, 2, \dots, n$ .

## 5.2 O PROBLEMA DE VALOR PRÓPRIO INVERSO MULTIPLICATIVO

Definição:

Dados uma matriz  $\mathbf{A}$  de dimensão  $n \times n$  real e simétrica e números reais  $\lambda_1^* \leq \lambda_2^* \leq \dots \leq \lambda_i^* \leq \dots \leq \lambda_n^*$ , encontrar um vetor  $\mathbf{c} \in \mathbb{R}^n$  de tal forma que os valores

próprios  $\lambda_1(\mathbf{c}) \leq \lambda_2(\mathbf{c}) \leq \dots \leq \lambda_i(\mathbf{c}) \leq \dots \leq \lambda_n(\mathbf{c})$  da matriz  $\mathbf{A}(\mathbf{c}) = \mathbf{D}\mathbf{A}$ , com  $\mathbf{D} = \text{diag}\{c_1, c_2, \dots, c_i, \dots, c_n\}$ , satisfaçam a igualdade  $\lambda_i(\mathbf{c}) = \lambda_i^*$  para  $i = 1, 2, \dots, n$ .

As considerações de simetria nas proposições acima não são essenciais, pois, em geral, um problema de valor próprio inverso é definido no campo dos números complexos (FRIEDLAND, 1977, p.15; BIEGLER-KÖNIG, 1981a, p.349). Para matrizes mais gerais onde valores próprios complexos aparecem, sempre há um esforço extra para ordená-los (CHEN; CHU, 1996, p. 2417).

Geralmente, um problema multiplicativo de valor próprio inverso pode ser escrito na forma de um problema aditivo de valor próprio inverso, fazendo  $\mathbf{A}_0 = \mathbf{0}$  e  $\mathbf{A}_k = \mathbf{e}_k \mathbf{a}_k^t$  para  $k = 1, \dots, n$  com  $\mathbf{a}_k^t$  sendo a  $k$ -ésima linha da matriz  $\mathbf{A}$  e  $\mathbf{e}_k$  o  $k$ -ésimo vetor unitário.

FRIEDLAND (1977, p.31) mostra que o problema multiplicativo de valor próprio inverso é sempre solucionável sobre o campo dos números complexos e que o número de soluções encontradas é no máximo igual a  $n!$  ( $n$  fatorial, sendo  $n$  a dimensão da matriz  $\mathbf{A}$ ). Porém quando restrito ao campo dos números reais não há garantia de solução, nem para o problema aditivo de valor próprio inverso nem para o problema multiplicativo de valor próprio inverso (CHEN; CHU, 1996, p.2417). As condições necessárias e suficientes para a existência da solução de um PVPI podem ser encontradas, por exemplo, nos trabalhos de BIEGLER-KÖNIG (1981B), LI (1991), LI (1995), SUN e YE (1986), SUN (1986) e XU (1992) entre outros.

Muitas aplicações encontradas do problema de valor próprio inverso, necessitam modificações na formulação apresentada em (5.1), tais como, o número de valores próprios pretendidos é menor que  $n$ , isto é, menor que a ordem da matriz; o número de elementos do vetor  $\mathbf{c}$  é diferente de  $n$ , isto é, pode ser maior, menor ou igual a  $n$ ; podem existir restrições sobre o vetor  $\mathbf{c}$ , por exemplo, pode-se exigir que  $\mathbf{c} \geq \mathbf{0}$ ; a existência de um modelo funcional que deve ser minimizado com restrições sobre os valores próprios de desigualdade em vez de igualdade.

Algumas aplicações onde o problema aditivo de valor próprio inverso aparece pode ser encontrada em FRIEDLAND et al. (1987, p. 635), entre elas cita-se, a solução do problema inverso de Sturm-Liouville, que tem origem no problema de

valor de contorno. Aplicações na espectroscopia nuclear, na espectroscopia molecular, no problema de comunalidade que tem origem na análise fatorial entre outras.

### 5.3 O PROBLEMA DE MÍNIMOS QUADRADOS DE VALOR PRÓPRIO INVERSO

Nas formulações práticas de um problema aditivo de valor próprio inverso freqüentemente ocorre que o número de parâmetros difere do número de valores próprios. Neste sentido considera-se a formulação de mínimos quadrados dada por,

$$\min_{\mathbf{c} \in \mathbb{R}^n} \sum_{i=1}^m (\lambda_i(\mathbf{c}) - \lambda_i^*)^2 \quad (5.2)$$

Reescrevendo a definição, numa formulação de mínimos quadrados para o problema aditivo de valor próprio inverso temos o seguinte:

Dadas, as matrizes  $n \times n$  reais e simétricas  $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_k, \dots, \mathbf{A}_\ell$  e os números reais  $\lambda_1^* \leq \lambda_2^* \leq \dots \leq \lambda_m^*$ , com  $m \leq n$ , encontrar o vetor  $\mathbf{c} \in \mathbb{R}^\ell$  de tal forma que a função

$$f(\mathbf{c}) = \frac{1}{2} \sum_{i=1}^m (\lambda_i(\mathbf{c}) - \lambda_i^*)^2 \quad (5.3)$$

seja minimizada. Na relação (5.3)  $\lambda_i(\mathbf{c})$  são os valores próprios de

$$\mathbf{A}(\mathbf{c}) = \mathbf{A}_0 + \sum_{k=1}^{\ell} c_k \mathbf{A}_k.$$

### 5.4 PROBLEMA DE VALOR PRÓPRIO INVERSO APLICADO NA OTIMIZAÇÃO DE PESOS EM LEVANTAMENTOS GEODÉSICOS

Considere-se a matriz dos coeficientes das equações normais

$$\mathbf{N} = \mathbf{A}^t \mathbf{P} \mathbf{A} \quad (5.4)$$

onde  $\mathbf{A}$ , é matriz de planejamento conhecida,  $\mathbf{P}$  é matriz diagonal tomada como incógnita e para este caso  $n$  é o número de observações e  $u$  número de parâmetros associados ao método dos mínimos quadrados na forma paramétrica.

A matriz  $\mathbf{N}$  assim definida é função dos elementos diagonais da matriz  $\mathbf{P}$ .

Exemplo:

$$\text{sejam as matrizes } {}_3\mathbf{A}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \text{ e } {}_3\mathbf{P}_3 = \begin{bmatrix} p_1 & & \\ & p_2 & \\ & & p_3 \end{bmatrix}.$$

Substituindo essas matrizes em  $\mathbf{N}$  resulta

$$\mathbf{N}(\mathbf{p}) = \begin{bmatrix} p_1 + p_3 & -p_3 \\ -p_3 & p_2 + p_3 \end{bmatrix}. \quad (5.5)$$

A matriz  $\mathbf{N}$  é função dos elementos diagonais de  $\mathbf{P}$  e escreve-se  $\mathbf{N}(\mathbf{p})$ . As modificações nos elementos diagonais de  $\mathbf{P}$  agem diretamente sobre a magnitude dos valores próprios da matriz  $\mathbf{N}$  e como consequência também em  $\mathbf{Q}_x = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+$ , os quais são usados na análise da qualidade de redes geodésicas. Com isso podem ser formuladas condições para tais valores próprios e determinar os pesos que levam a (5.5) a apresentar características ótimas em termos de valores próprios.

Especificamente no caso da otimização dos pesos das observações geodésicas e considerando a matriz dos pesos na forma diagonal, a expressão (5.4) é escrita na forma (5.1), como

$$\mathbf{N} = \sum_{j=1}^n p_j \mathbf{a}_j^t \mathbf{a}_j \quad (5.6)$$

onde  $n$  é o numero de observações,  $\mathbf{a}_j$  é a  $j$ -ésima linha da matriz planejamento  $\mathbf{A}$  e  $p_j = p_{jj}$  é cada elemento da diagonal principal da matriz dos pesos  $\mathbf{P}$  e representa o peso atribuído a  $j$ -ésima observação. Desta forma é estabelecido a ligação entre a otimização dos pesos das observações geodésicas e um problema de valor próprio inverso.

## 5.5 FORMULAÇÃO NUMÉRICA

Sobre a formulação numérica de um problema de valor próprio inverso, classicamente, há duas formas de caracterizar este problema quando se considera as grandezas  $\lambda_1(\mathbf{c}) \leq \lambda_2(\mathbf{c}) \leq \dots \leq \lambda_i(\mathbf{c}) \leq \dots \leq \lambda_n(\mathbf{c})$  como sendo os valores próprios da

matriz  $\mathbf{A}(\mathbf{c}) = \mathbf{A}_0 + \sum_{k=1}^n c_k \mathbf{A}_k$ .

### 5.5.1 Formulação I

Dados  $n$  números reais  $\lambda_1^* \leq \lambda_2^* \leq \dots \leq \lambda_i^* \leq \dots \leq \lambda_n^*$ , encontrar o vetor  $\mathbf{c}^* \in \mathbb{R}^n$ , tal que  $\lambda_i(\mathbf{c}^*) = \lambda_i^*$  para  $i = 1, 2, \dots, n$ . Esta formulação pode ser colocada como um problema de solucionar o sistema não-linear,  $\mathbf{F}(\mathbf{c}) = \mathbf{0}$  onde

$$\mathbf{F}(\mathbf{c}) = \begin{bmatrix} \lambda_1(\mathbf{c}) - \lambda_1^* \\ \vdots \\ \lambda_i(\mathbf{c}) - \lambda_i^* \\ \vdots \\ \lambda_n(\mathbf{c}) - \lambda_n^* \end{bmatrix} \quad (5.7)$$

com  $\lambda_1(\mathbf{c}) \leq \lambda_2(\mathbf{c}) \leq \dots \leq \lambda_i(\mathbf{c}) \leq \dots \leq \lambda_n(\mathbf{c})$ .

### 5.5.2 Formulação II

Uma segunda formulação pode ser enunciada da seguinte forma:

Seja

$$\mathbf{A}(\mathbf{c}) - \lambda_i^* \mathbf{I}, \text{ com } i = 1, 2, \dots, n. \quad (5.8)$$

Se  $\lambda_i^*$  for um valor próprio de  $\mathbf{A}(\mathbf{c})$ , então  $\det(\mathbf{A}(\mathbf{c}) - \lambda_i^* \mathbf{I}) = 0$  para  $i = 1, 2, \dots, n$ .

Para esta formulação resolve-se o sistema de equações não-lineares:

$$\mathbf{G}(\mathbf{c}) = \begin{bmatrix} \det(\mathbf{A}(\mathbf{c}) - \lambda_1^* \mathbf{I}) \\ \vdots \\ \det(\mathbf{A}(\mathbf{c}) - \lambda_i^* \mathbf{I}) \\ \vdots \\ \det(\mathbf{A}(\mathbf{c}) - \lambda_n^* \mathbf{I}) \end{bmatrix} = \mathbf{0} \quad (5.9)$$

ou de forma equivalente, utilizando a propriedade (h) dos valores próprios,

$$\mathbf{G}(\mathbf{c}) = \begin{bmatrix} \prod_{k=1}^n (\lambda_k(\mathbf{c}) - \lambda_1^*) \\ \vdots \\ \prod_{k=1}^n (\lambda_k(\mathbf{c}) - \lambda_i^*) \\ \vdots \\ \prod_{k=1}^n (\lambda_k(\mathbf{c}) - \lambda_n^*) \end{bmatrix} = \mathbf{0} \quad (5.10)$$

onde  $\prod_{k=1}^n$  indica o produtório de 1 a n.

## 6 O PROBLEMA NÃO LINEAR

As formulações (5.7) e (5.9) são sistemas de equações não-lineares. Algumas técnicas de otimização numérica podem ser aplicadas com o intuito de encontrar a solução. Em particular, como será visto mais a diante, o método de Newton será utilizado para solucionar  $\nabla f = \mathbf{0}$ .

Um Sistema de Equações Não-Lineares pode ser caracterizado pela seguinte definição:

Dada  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,

Encontrar  $\mathbf{x}^* \in \mathbb{R}^n$  para o qual se tem  $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$ , (6.1)

A forma (6.1) é a forma padrão para representar um sistema de n equações a n incógnitas.

Exemplo:

$$\mathbf{F}(x_1, x_2) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix}$$

em que uma de suas soluções é  $\mathbf{x}^* = [0; 3]^t$ , ou seja,  $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$ .

O valor  $\mathbf{x}^*$  que soluciona (6.1) é o minimizador da função,

$$f(\mathbf{x}) = \sum_{i=1}^n [F_i(\mathbf{x})]^2 \tag{6.2}$$

onde  $F_i(\mathbf{x})$  representa a i-ésima componente da função  $\mathbf{F}$ .

A situação descrita acima caracteriza um problema de minimização irrestrita, definido como:

Dada  $f : \mathbb{R}^n \rightarrow \mathbb{R}$

Encontrar  $\mathbf{x}^* \in \mathbb{R}^n$  para o qual se obtém  $f(\mathbf{x}^*) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n$ . (6.3)

Tanto do ponto de vista do “desenho” algorítmico quanto da possibilidade da solução a formulação descrita acima apresenta consideráveis vantagens como poderá ser observado posteriormente.

Exemplo:

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(x_1, x_2) = (x_1 - 2)^2 + (x_2 + 8)^4$$

cuja solução é  $\mathbf{x}^* = [2, -8]^t$ .

A função  $f$  como construída acima pode ser interpretada como uma função obtida a partir de um sistema de equações não-lineares  $\mathbf{F}$ ,

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(x_1, x_2) = \begin{bmatrix} (x_1 - 2) \\ (x_2 + 8)^2 \end{bmatrix}$$

e aplicando-se (6.2) se obtém

$$f(\mathbf{x}) = f(x_1, x_2) = (x_1 - 2)^2 + (x_2 + 8)^4.$$

Uma possibilidade para encontrar  $\mathbf{x}^*$  para o qual  $\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$ , é minimizar  $f(\mathbf{x})$ , problema este que pode apresentar vantagens quanto a garantias de solubilidade.

Uma solução  $\mathbf{x}^*$  onde a função atinge seu menor valor é dita ser um minimizador global de  $f$  e escreve-se:

Um ponto  $\mathbf{x}^*$  é um minimizador global de  $f$  se  $f(\mathbf{x}^*) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathbb{R}^n$  ou sobre um domínio de interesse.

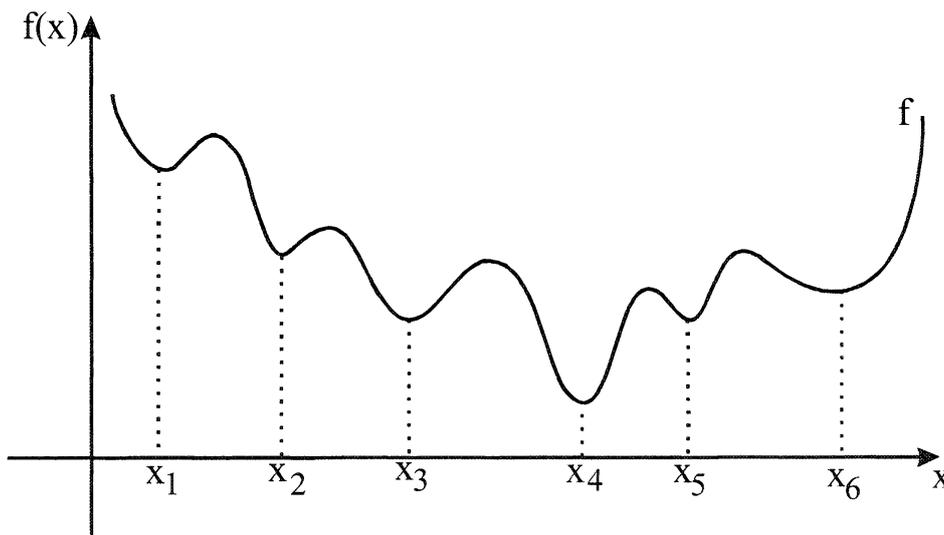
Um minimizador global pode ser difícil de ser encontrado, em face que o conhecimento que se tem da função é somente local (NOCEDAL; WRIGHT, 1999, p.13). A maioria dos algoritmos é capaz de encontrar somente um minimizador local, que é um ponto em que a função obtém o menor valor em uma determinada vizinhança. Uma definição é:

Um ponto  $\mathbf{x}^*$  é um minimizador local se existir uma região aberta contendo  $\mathbf{x}^*$ , representada por  $\mathfrak{N}$  tal que  $f(\mathbf{x}^*) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathfrak{N}$ .

Um ponto que satisfaz esta definição é denominado de minimizador local fraco. Esta terminologia distingue-o de um minimizador local estrito, enunciado como segue:

Um ponto  $\mathbf{x}^*$  é um minimizador local estrito se existe uma região aberta contendo  $\mathbf{x}^*$ , representada por  $\mathfrak{N}$ , tal que  $f(\mathbf{x}^*) < f(\mathbf{x})$  para todo  $\mathbf{x} \in \mathfrak{N}$  com  $\mathbf{x} \neq \mathbf{x}^*$ .

FIGURA 6.1- MINIMIZADOR LOCAL E GLOBAL



A figura 6.1 ilustra uma função com vários minimizadores locais. O minimizador global representado por  $x_4$  é usualmente difícil de ser encontrado de acordo com DENNIS JR. e SCHNABEL (1983, p.5), visto que, os algoritmos tendem a estacionar em minimizadores locais. De forma geral, usa-se o termo global, no sentido que o método é global ou então que o algoritmo é globalmente convergente, denotando que o método utilizado converge para um mínimo local de uma função não-linear ou então de um sistema de equações não-lineares, independente do ponto inicial utilizado. Às vezes, tendo algumas informações adicionais sobre  $f$ , pode-se identificar um minimizador global. Um caso particular é quando a função  $f$  é convexa<sup>3</sup>, na qual todo minimizador local é também um minimizador global.

Se  $f$  é continuamente diferenciável, pode-se reconhecer um mínimo local examinando o gradiente de  $f$ ,  $\nabla f(\mathbf{x}^*)$ , e a matriz Hessiana de  $f$ ,  $\nabla^2 f(\mathbf{x}^*)$ . Para isso são enunciados os seguintes teoremas, cujas provas podem ser encontradas em NOCEDAL e WRIGHT (1999, p. 15):

TEOREMA 6.1 (condição necessária de primeira ordem)

Se  $\mathbf{x}^*$  é um minimizador local e  $f$  é continuamente diferenciável em uma região aberta contendo  $\mathbf{x}^*$ , então  $\nabla f(\mathbf{x}^*) = \mathbf{0}$ .

TEOREMA 6.2 (condição necessária de segunda ordem)

<sup>3</sup> Sobre convexidade ver, por exemplo, os autores: FRIEDLANDER (1994) e MARTINEZ e SANTOS (1995).

Se  $\mathbf{x}^*$  é um minimizador local de  $f$  e  $\nabla^2 f$  é continuamente diferenciável em uma região aberta contendo  $\mathbf{x}^*$  então  $\nabla f(\mathbf{x}^*) = \mathbf{0}$  e  $\nabla^2 f(\mathbf{x}^*)$  é semidefinida positiva.

TEOREMA 6.3 (condição suficiente de segunda ordem)

Supondo que  $\nabla^2 f$  é continuamente diferenciável em uma região aberta contendo  $\mathbf{x}^*$  e que  $\nabla f(\mathbf{x}^*) = \mathbf{0}$  e  $\nabla^2 f(\mathbf{x}^*)$  é definida positiva. Então  $\mathbf{x}^*$  é um minimizador estrito de  $f$ .

TEOREMA 6.4 (convexidade)

Quando  $f$  é convexa, qualquer minimizador local  $\mathbf{x}^*$  é um minimizador global de  $f$ . Se ainda  $f$  é diferenciável, então qualquer ponto estacionário  $\mathbf{x}^*$  é minimizador global de  $f$ .

O teorema 6.1 e o teorema 6.2 referem-se às condições necessárias de otimalidade. O teorema 6.3 refere-se à condição suficiente de otimalidade e o teorema 6.4 caracteriza minimizadores locais e minimizador global quando a função objetivo é convexa.

Estes resultados fundamentam os algoritmos de minimização irrestrita. De uma forma ou outra, todos os algoritmos buscam um ponto onde o gradiente de  $f$  se anula.

## 6.1 ALGORITMOS

Otimização é um problema de minimizar ou maximizar uma função, que às vezes pode ser, sujeita a restrições em suas variáveis. Os algoritmos de otimização são, em geral, iterativos. Iniciam com um valor arbitrário gerando uma seqüência de estimativas melhoradas em busca da solução.

Em geral os algoritmos requerem um valor inicial arbitrário  $\mathbf{x}^0$  da solução para sua inicialização. Quando se tem informações suficientes sobre o problema, pode-se dar um valor inicial aproximado adequado para  $\mathbf{x}^0$ . Isto é,  $\mathbf{x}^0$  pode ser uma

boa estimativa para a solução do problema. Nas situações práticas, em regra, não se tem informação se o valor inicial  $\mathbf{x}^0$  é suficientemente adequado e na maioria das vezes é escolhido de forma arbitrária.

Iniciando com um valor arbitrário  $\mathbf{x}^0$ , os algoritmos de otimização geram uma seqüência de iterações, denotada por  $\{\mathbf{x}^k\}_{k=0}^{\infty}$ , que só termina quando, não se pode mais fazer progresso com o algoritmo, ou ainda se pode aproximar da solução tanto quanto se queira, isto é, com uma precisão suficiente (convergência).

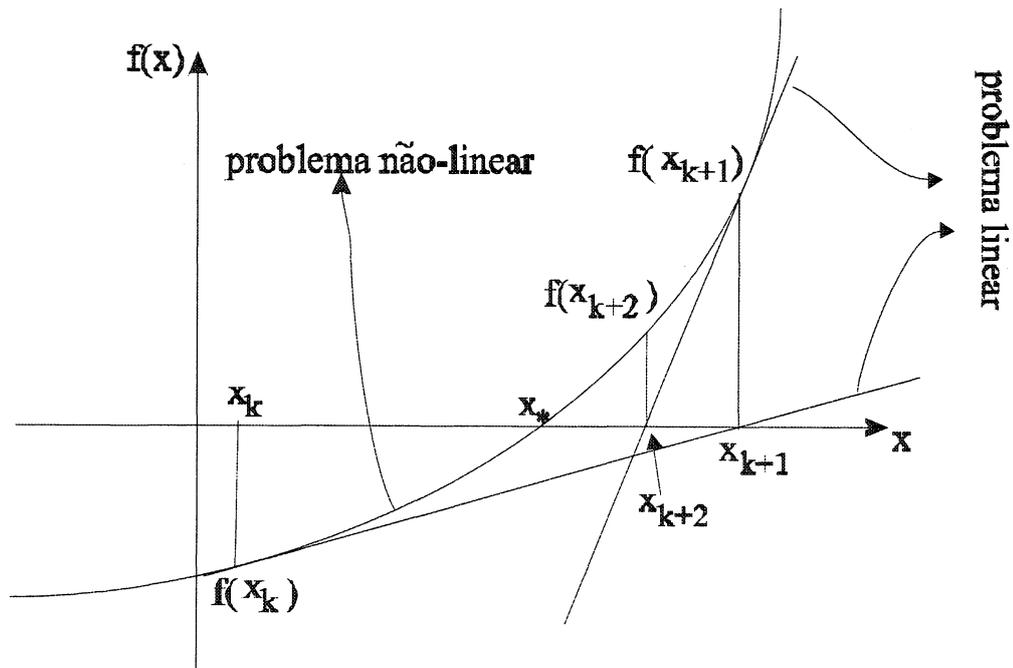
Para decidir como mover-se de uma iteração, por exemplo,  $\mathbf{x}^k$  para a próxima  $\mathbf{x}^{k+1}$  o algoritmo usa, em geral, informações sobre a função no ponto  $\mathbf{x}^k$ . Esta informação é usada para encontrar o novo valor aproximado melhorado  $\mathbf{x}^{k+1}$  cujo valor funcional, para algoritmos de minimização, deve ser menor que no ponto  $\mathbf{x}^k$ .

Basicamente duas estratégias podem ser usadas para mover-se de  $\mathbf{x}^k$  para a nova iteração  $\mathbf{x}^{k+1}$ , com decréscimo do valor da função, que são as estratégias de busca linear e de região de confiança tratadas nos próximos itens.

### 6.1.1 O Método de Newton

O princípio do método de Newton consiste em considerar, que a solução de  $\mathbf{F}(\mathbf{x})$  em um passo é um problema não-linear. Então a solução para  $\mathbf{F}(\mathbf{x}) = \mathbf{0}$  irá se aproximando através de uma seqüência de pontos  $\{\mathbf{x}^k\}$ . Em cada aproximação  $\mathbf{x}^k$ , constrói-se, com as informações disponíveis neste ponto, um problema linear, que teoricamente, a solução pode ser encontrada. O valor atualizado  $\mathbf{x}^{k+1}$  será a solução do problema linear, o qual muda de uma iteração para a outra, e em caso de convergência, estes valores atualizados estarão cada vez mais próximos da solução do problema original que é não-linear, ver figura 6.2.

FIGURA 6.2 – MÉTODO DE NEWTON: CONVERGÊNCIA



Considerando o desenvolvimento da série de Taylor até a derivada primeira de  $F(\mathbf{x})$ , em uma vizinhança do ponto atual  $\mathbf{x}^k$ , se obtém a aproximação linear para  $F(\mathbf{x})$  no ponto  $\mathbf{x}^k$

$$F(\mathbf{x}) \approx L^k(\mathbf{x}) = F(\mathbf{x}^k) + \mathbf{J}(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k) \quad (6.4)$$

onde

$$\mathbf{J}(\mathbf{x}^k) = \left[ \frac{\partial F_i(\mathbf{x}^k)}{\partial x_j} \right]_{\substack{i=1,2,\dots,m \\ j=1,2,\dots,n}} \quad (6.5)$$

é a matriz Jacobiana de  $F(\mathbf{x})$ , obtida avaliando-se as derivadas primeiras das funções  $F_i(\mathbf{x})$   $i = 1, 2, \dots, m$ .

O ponto  $\mathbf{x}^{k+1}$  é obtido solucionando  $L^k(\mathbf{x}) = \mathbf{0}$ , ou de forma equivalente,

$$F(\mathbf{x}^k) + \mathbf{J}(\mathbf{x}^k)(\mathbf{x} - \mathbf{x}^k) = \mathbf{0} \quad (6.6)$$

Uma iteração no método de Newton consiste em resolver o sistema linear da equação (6.6) da seguinte forma,

$$\begin{aligned} \mathbf{J}(\mathbf{x}^k) \mathbf{s}_k &= -\mathbf{F}(\mathbf{x}^k) \\ \mathbf{x}^{k+1} &= \mathbf{x}^k + \mathbf{s}_k \end{aligned} \quad (6.7)$$

onde  $\mathbf{s}_k = \mathbf{x}^{k+1} - \mathbf{x}^k$ .

O método de Newton é na sua forma básica, um método iterativo local, no sentido em que se pode garantir apenas a convergência à solução, supondo que o ponto usado como valor inicial já seja suficientemente próximo da solução. Em muitos casos, os métodos locais convergem mesmo se a aproximação inicial não for “muito boa” (MARTINEZ; SANTOS, 1995, p.74). Um exemplo deste fato é quando um sistema linear é solucionado como se fosse não-linear, desta forma o método de Newton obtém uma solução independentemente da escolha do ponto inicial.

Por isso considera-se, até certo ponto que seja natural o uso do método de Newton, na sua forma básica, pela praticidade de sua aplicação e as razoáveis possibilidades de sucesso.

A aplicação do método de Newton na solução de um problema de valor próprio inverso tem seu desenvolvimento baseado nas seguintes considerações:

#### 6.1.1.1 O Método de Newton para solução da formulação I

A partir da relação de normalidade tem-se

$$\mathbf{m}_i^t(\mathbf{c}) \mathbf{m}_i(\mathbf{c}) = 1, \quad (6.8)$$

em que  $\mathbf{m}_i(\mathbf{c})$  é o vetor próprio, em função de  $\mathbf{c}$ , correspondente ao valor próprio  $\lambda_i(\mathbf{c})$ .

Multiplicando (6.8) por  $\lambda_i(\mathbf{c})$  tem-se

$$\mathbf{m}_i^t(\mathbf{c}) \lambda_i(\mathbf{c}) \mathbf{m}_i(\mathbf{c}) = \lambda_i(\mathbf{c}). \quad (6.9)$$

considerando que  $\lambda_i \mathbf{m}_i(\mathbf{c}) = \mathbf{A}(\mathbf{c}) \mathbf{m}_i(\mathbf{c})$  e diferenciando a relação (6.9), obtém-se a variação do valor próprio  $\lambda_i$  em relação  $c_j$ , dada por

$$\frac{\partial \lambda_i(\mathbf{c})}{\partial c_j} = \mathbf{m}_i^t(\mathbf{c}) \mathbf{A}_j \mathbf{m}_i(\mathbf{c}). \quad (6.10)$$

O cálculo da segunda derivada  $\frac{\partial^2 \lambda_i}{\partial c_j \partial c_p} = \lambda_{i,j,p}$   $j, p = 1, 2, \dots, n$  é obtida após

a diferenciação total de (6.10),

$$\lambda_{i,j,p} = 2 \sum_{\substack{s=1 \\ s \neq i}}^n \frac{(\mathbf{m}_s^t \mathbf{A}_p \mathbf{m}_i)(\mathbf{m}_s^t \mathbf{A}_j \mathbf{m}_i)}{\lambda_i - \lambda_s} \quad \text{para } \lambda_i \neq \lambda_s \quad (6.11)$$

onde  $\frac{\partial \mathbf{A}(\mathbf{c})}{\partial c_j} = \mathbf{A}_j$  e  $\frac{\partial \mathbf{A}(\mathbf{c})}{\partial c_p} = \mathbf{A}_p$ .

Para mais informações sobre este tópico consultar OLIVEIRA (2003).

Com as considerações acima a matriz Jacobiana é escrita como

$$\mathbf{J}_{i,j}(\mathbf{c}) = \mathbf{m}_i^t(\mathbf{c}) \mathbf{A}_j \mathbf{m}_i(\mathbf{c}). \quad (6.12)$$

Os passos para solucionar um problema de valor próprio inverso através do método de Newton, utilizando a formulação I, são:

escolher um valor inicial aproximado  $\mathbf{c}^0$  e para  $k = 0, 1, 2, \dots$  fazer

1. Formar a matriz  $\mathbf{A}(\mathbf{c}^k)$ ;
2. obter os valores próprios  $\lambda(\mathbf{c}^k)$  e os vetores próprios  $\mathbf{m}(\mathbf{c}^k)$  de  $\mathbf{A}(\mathbf{c}^k)$ ;
3. parar se  $\|\lambda(\mathbf{c}^k) - \lambda^*\| < \varepsilon$ , onde  $\varepsilon$  representa a qualidade da aproximação, dado em um critério de parada estabelecido e  $\|(\bullet)\|$  é uma norma qualquer de um vetor;
4. formar a matriz jacobiana  $\mathbf{J}(\mathbf{c}^k)$ ;

5. obter  $\mathbf{c}^{k+1}$  solucionando o sistema  $\mathbf{J}(\mathbf{c}^k)(\mathbf{c}^{k+1} - \mathbf{c}^k) = -\mathbf{F}(\mathbf{c}^k)$  e voltar ao passo 1.

Com a finalidade de interromper o processo iterativo, considerando que houve divergência, pode-se usar o critério que, para algum  $k$ , a norma  $\|\lambda(\mathbf{c}^k) - \lambda^*\|$  é maior que uma determinada tolerância permitida. Por exemplo,  $\|\lambda(\mathbf{c}^k) - \lambda^*\| > 10^8$ .

### 6.1.1.2 O método de Newton para solução da formulação II

A  $i$ -ésima equação de (5.10) é

$$g_i(\mathbf{c}) = \prod_{k=1}^n (\lambda_k(\mathbf{c}) - \lambda_i^*) \quad (6.13)$$

desenvolvendo (6.13) se obtém

$$g_i(\mathbf{c}) = (\lambda_1(\mathbf{c}) - \lambda_i^*) \times (\lambda_2(\mathbf{c}) - \lambda_i^*) \times \dots \times (\lambda_n(\mathbf{c}) - \lambda_i^*). \quad (6.14)$$

Com a equação (6.14) considerando também

$$\lambda_i(\mathbf{c}) = \mathbf{m}_i^t(\mathbf{c})\mathbf{A}(\mathbf{c})\mathbf{m}_i(\mathbf{c}) \quad (6.15)$$

e

$$\frac{\partial \lambda_i(\mathbf{c})}{\partial c_j} = \mathbf{m}_i^t(\mathbf{c}) \frac{\partial \mathbf{A}(\mathbf{c})}{\partial c_j} \mathbf{m}_i(\mathbf{c}) \quad (6.16)$$

com  $\frac{\partial \lambda_i(\mathbf{c})}{\partial c_j} = \lambda_{i,j}$  e  $\frac{\partial \mathbf{A}(\mathbf{c})}{\partial c_j} = \mathbf{A}_j$  a (6.16) torna-se

$$\lambda_{i,j} = \mathbf{m}_i^t(\mathbf{c})\mathbf{A}_j\mathbf{m}_i(\mathbf{c}). \quad (6.17)$$

Com isto, a variação diferencial da equação (6.14) em relação à  $c_j$  é

$$\frac{\partial g_i(\mathbf{c})}{\partial c_j} = \sum_{k=1}^n \left[ \mathbf{m}_k^t(\mathbf{c})\mathbf{A}_j\mathbf{m}_k(\mathbf{c}) \prod_{\substack{r=1 \\ r \neq k}}^n (\lambda_r(\mathbf{c}) - \lambda_i^*) \right] \quad (6.18)$$

Com a finalidade de utilizar o método de Newton se faz necessário obter a matriz Jacobiana da (5.10), a qual pode ser calculada com o uso da equação (6.18). Desta forma a matriz Jacobiana de  $g(\mathbf{c})$ , representada por  $\mathbf{G}(\mathbf{c})$  é

$$\mathbf{G}(\mathbf{c}) = \begin{bmatrix} \sum_{k=1}^n \left[ \mathbf{m}_k^t(\mathbf{c}) \mathbf{A}_1 \mathbf{m}_k(\mathbf{c}) \prod_{\substack{r=1 \\ r \neq k}}^n (\lambda_r(\mathbf{c}) - \lambda_i^*) \right] & \cdots & \sum_{k=1}^n \left[ \mathbf{m}_k^t(\mathbf{c}) \mathbf{A}_n \mathbf{m}_k(\mathbf{c}) \prod_{\substack{r=1 \\ r \neq k}}^n (\lambda_r(\mathbf{c}) - \lambda_i^*) \right] \\ \vdots & & \vdots \\ \sum_{k=1}^n \left[ \mathbf{m}_k^t(\mathbf{c}) \mathbf{A}_1 \mathbf{m}_k(\mathbf{c}) \prod_{\substack{r=1 \\ r \neq k}}^n (\lambda_r(\mathbf{c}) - \lambda_i^*) \right] & \cdots & \sum_{k=1}^n \left[ \mathbf{m}_k^t(\mathbf{c}) \mathbf{A}_n \mathbf{m}_k(\mathbf{c}) \prod_{\substack{r=1 \\ r \neq k}}^n (\lambda_r(\mathbf{c}) - \lambda_i^*) \right] \end{bmatrix} \quad (6.19)$$

ou de forma equivalente,

$$\mathbf{G}(\mathbf{c}) = \begin{bmatrix} \prod_{r \neq 1} (\lambda_r(\mathbf{c}) - \lambda_1^*) & \prod_{r \neq 2} (\lambda_r(\mathbf{c}) - \lambda_1^*) & \cdots & \prod_{r \neq n} (\lambda_r(\mathbf{c}) - \lambda_1^*) \\ \vdots & \vdots & \cdots & \vdots \\ \prod_{r \neq 1} (\lambda_r(\mathbf{c}) - \lambda_n^*) & \prod_{r \neq 2} (\lambda_r(\mathbf{c}) - \lambda_n^*) & \cdots & \prod_{r \neq n} (\lambda_r(\mathbf{c}) - \lambda_n^*) \end{bmatrix} \times \begin{bmatrix} \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_1 \mathbf{m}_1(\mathbf{c}) & \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_2 \mathbf{m}_1(\mathbf{c}) & \cdots & \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_n \mathbf{m}_1(\mathbf{c}) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_1 \mathbf{m}_1(\mathbf{c}) & \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_2 \mathbf{m}_1(\mathbf{c}) & \cdots & \mathbf{m}_1^t(\mathbf{c}) \mathbf{A}_n \mathbf{m}_1(\mathbf{c}) \end{bmatrix} \quad (6.20)$$

A (6.20) escrita de uma forma compacta torna-se,

$$\mathbf{G}(\mathbf{c}) = \text{diag} \left( \frac{g_i(\mathbf{c})}{f_i(\mathbf{c})} \right) \mathbf{V}(\mathbf{c}) \quad (6.21)$$

com  $\mathbf{V}(\mathbf{c})$  definida por

$$\mathbf{V}_{ij}(\mathbf{c}) = (\lambda_i(\mathbf{c}) - \lambda_i^*) \sum_{k=1}^n \frac{\mathbf{m}_k^t(\mathbf{c}) \mathbf{A}_j \mathbf{m}_k(\mathbf{c})}{\lambda_k(\mathbf{c}) - \lambda_i^*} \quad (6.22)$$

A Formulação I é mais simples ou mais natural, e foi usada por DOWNING e HOUSEHOLDER (1956) no caso do problema aditivo e também por KUBLANOVSKAJA (1970) para o caso geral. A Formulação II foi proposta por BIEGLER-KÖNIG (1981a) e apresenta algumas desvantagens em relação à Formulação I, como pode ser visto em NOCEDAL e OVERTON (1983).

Abaixo é apresentado um exemplo sugestionando que a formulação I apresenta vantagens em relação à formulação II,

### Exemplo

Considera-se para este exemplo o caso de dimensão  $2 \times 2$  onde as matrizes  $\mathbf{A}_i$  são matrizes diagonais dadas por,

$$\mathbf{A}_1 = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix} \text{ e } \mathbf{A}_2 = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

utilizando a expressão (5.1) se obtém

$$\mathbf{A}(\mathbf{c}) = \begin{bmatrix} c_1 a_1 + c_2 b_1 & 0 \\ 0 & c_1 a_2 + c_2 b_2 \end{bmatrix}$$

e lembrando que os valores próprios de uma matriz diagonal coincidem com os elementos da diagonal principal consegue-se  $\lambda_1(\mathbf{c}) = c_1 a_1 + c_2 b_1$  e  $\lambda_2(\mathbf{c}) = c_1 a_2 + c_2 b_2$ . Utilizando a formulação I para solucionar o problema de valor próprio inverso tem-se o sistema linear,

$$\begin{cases} c_1 a_1 + c_2 b_1 - \lambda_1^* = 0 \\ c_1 a_2 + c_2 b_2 - \lambda_2^* = 0 \end{cases} \quad (6.23)$$

De outra forma se for utilizada a formulação II calcula-se dois determinantes, um para cada equação, ou seja,

$$\det(\mathbf{A}(\mathbf{c}) - \lambda_1^* \mathbf{I}) = (c_1 a_1 + c_2 b_1 - \lambda_1^*)(c_1 a_2 + c_2 b_2 - \lambda_1^*)$$

$$\det(\mathbf{A}(\mathbf{c}) - \lambda_2^* \mathbf{I}) = (c_1 a_1 + c_2 b_1 - \lambda_2^*)(c_1 a_2 + c_2 b_2 - \lambda_2^*)$$

que dá origem ao sistema de equações não-lineares (duas equações quadráticas),

$$\begin{cases} (c_1 a_1 + c_2 b_1 - \lambda_1^*)(c_1 a_2 + c_2 b_2 - \lambda_1^*) = 0 \\ (c_1 a_1 + c_2 b_1 - \lambda_2^*)(c_1 a_2 + c_2 b_2 - \lambda_2^*) = 0 \end{cases} \quad (6.24)$$

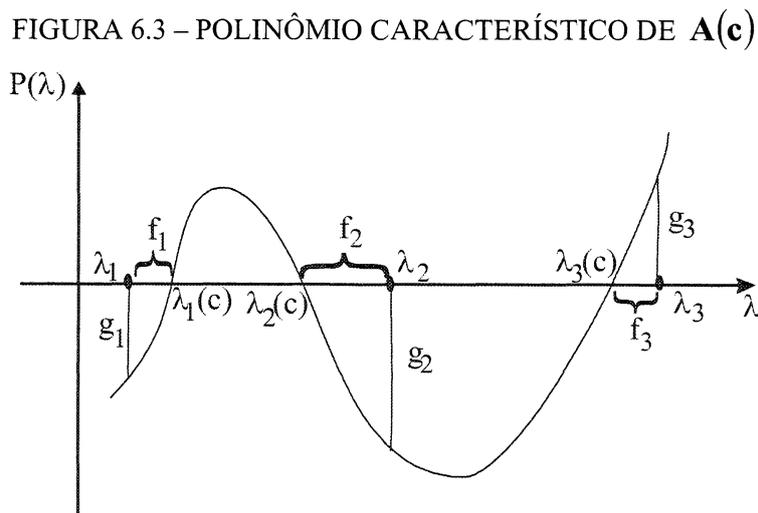
Se for aplicado, por exemplo, o método de Newton para solucionar ambos os sistemas consegue-se encontrar a solução de (6.23) em um único passo enquanto que, se a mesma estratégia é utilizada em (6.24) pode-se levar alguns passos para chegar à solução.

Como visto anteriormente a formulação II, em geral, oferece complicações ao problema. Observa-se reescrevendo a formulação (5.9) usando (5.7)

$$g_i(\mathbf{c}) = f_i(\mathbf{c}) \prod_{k \neq i} (\lambda_k(\mathbf{c}) - \lambda_i^*) \quad (6.25)$$

que cada equação do sistema (5.7) fica multiplicado por um polinômio de grau  $n - 1$ .

Uma comparação gráfica entre as Formulações I e II é apresentada na figura 6.3, na qual se considera o gráfico do polinômio característico  $P(\lambda)$  de  $\mathbf{A}(\mathbf{c})$  e os correspondentes valores próprios pretendidos  $\lambda_i^*$  ( $i = 1, 2, 3$ ).



Enquanto a Formulação I trata de minimizar  $f_1, f_2, \dots, f_n$  a Formulação II se refere à minimização de  $g_1, g_2, \dots, g_n$ .

## 6.2 OTIMIZAÇÃO: MINIMIZAÇÃO IRRESTRITA

Matematicamente, otimização é a minimização ou maximização de uma função sujeita a restrições em suas variáveis. Na minimização irrestrita, minimiza-se uma função objetivo, que depende de variáveis reais, sem restrições sobre estas variáveis. A formulação matemática, equivalente a (6.3) é

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad (6.26)$$

onde  $\mathbf{x} \in \mathbb{R}^n$  é um vetor com  $n \geq 1$  componentes.

O problema de equações não lineares pode ser representado como segue:

Dada a função  $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , encontrar  $\mathbf{x}^* \in \mathbb{R}^n$ , tal que

$$F(\mathbf{x}^*) = \mathbf{0} \quad (6.27)$$

Um passo de Newton para (6.27) é obtido como na relação (6.6). Porém, o passo como colocado em (6.6) não garante convergência global. Nestas condições para aceitar  $\mathbf{x}^{k+1}$  para a próxima iteração, é razoável ocorrer que  $\|F(\mathbf{x}^{k+1})\|$  seja menor que  $\|F(\mathbf{x}^k)\|$  para alguma norma  $\|\bullet\|$ . Se por conveniência a norma escolhida é a euclidiana escreve-se,

$$\|F(\mathbf{x})\|_2^2 = F(\mathbf{x})^t F(\mathbf{x}). \quad (6.28)$$

Se é requerido que em cada passo do processo iterativo  $\|F(\mathbf{x})\|_2$  diminua, formula-se uma função de minimização como,

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}): \frac{1}{2} F(\mathbf{x})^t F(\mathbf{x}) \quad (6.29)$$

Comparando a equação (6.27) com a formulação (6.29) se percebe que uma solução para a equação (6.27) coincide com uma solução para (6.29). Então, pode-se tentar solucionar a (6.27) minimizando a (6.29).

Uma estratégia de solução globalizada para a (6.27) é baseada sobre a estratégia de minimização do problema (6.29) fazendo uso de aproximações quadráticas à  $f(\mathbf{x})$ .

Neste caso o método de Newton usa de uma aproximação quadrática para (6.29) em cada iteração e obtém o minimizador desta quadrática.

A aproximação de  $f$  em um ponto  $\mathbf{x}^k$ , por meio de uma função quadrática, utilizando a expansão truncada de ordem 2 de Taylor, é representado por,

$$m_c(\mathbf{x}) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^t (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^k)^t \nabla^2 f(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) \quad (6.30)$$

onde,

$m_c(\mathbf{x})$  representa o modelo quadrático corrente ou modelo quadrático atual;

$f(\mathbf{x}^k) = \frac{1}{2} \mathbf{F}(\mathbf{x}^k)^t \mathbf{F}(\mathbf{x}^k)$  é o valor numérico da função (6.29) no ponto  $\mathbf{x}^k$ ;

$\nabla f(\mathbf{x}^k) = \mathbf{J}(\mathbf{x}^k)^t \mathbf{F}(\mathbf{x}^k)$  é a avaliação do gradiente da função  $f(\mathbf{x}^k)$  no ponto  $\mathbf{x}^k$  e;

$\nabla^2 f(\mathbf{x}^k) = \mathbf{J}(\mathbf{x}^k)^t \mathbf{J}(\mathbf{x}^k) + \mathbf{S}(\mathbf{x}^k)$  é a avaliação da matriz da segunda derivada da função (6.29) no ponto  $\mathbf{x}^k$ , com

$$\mathbf{S}(\mathbf{x}^k) = \sum_{j=1}^m (\mathbf{F}_j(\mathbf{x}^k) \nabla^2 \mathbf{F}_j(\mathbf{x}^k)) \quad (6.31)$$

apresentando as informações sobre a segunda derivada da função  $\mathbf{F}(\mathbf{x})$  no ponto  $\mathbf{x}^k$ .

A matriz  $\mathbf{J}(\mathbf{x})$  é a matriz Jacobiana de  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , como colocado na (6.5) e reescrita explicitamente como,

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mathbf{F}_1(\mathbf{x})}{\partial x_1} & \frac{\partial \mathbf{F}_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \mathbf{F}_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial \mathbf{F}_2(\mathbf{x})}{\partial x_1} & \frac{\partial \mathbf{F}_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \mathbf{F}_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \mathbf{F}_m(\mathbf{x})}{\partial x_1} & \frac{\partial \mathbf{F}_m(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \mathbf{F}_m(\mathbf{x})}{\partial x_n} \end{bmatrix}. \quad (6.32)$$

A matriz  $\nabla^2 \mathbf{F}_j(\mathbf{x})$  é obtida da seguinte forma,

$$\nabla^2 F_j(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 F_j(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F_j(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_2^2} & \dots & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F_j(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 F_j(\mathbf{x})}{\partial x_n^2} \end{bmatrix}. \quad (6.33)$$

Do cálculo diferencial tem-se o resultado

$$\frac{\partial F_j(\mathbf{x})}{\partial x_1 \partial x_2} = \frac{\partial F_j(\mathbf{x})}{\partial x_2 \partial x_1} \quad (6.34)$$

logo a matriz dada em (6.33) é necessariamente uma matriz simétrica.

Exemplo:

Considerando  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(x_1, x_2) = \begin{bmatrix} x_1^3 - x_2^2 - 3 \\ 3x_1^2 - x_2 + 4 \\ x_1^4 - x_2 \end{bmatrix}$$

e formulando o problema de minimização, como exposto na relação (6.29) se escreve

$$f(\mathbf{x}) = \frac{1}{2} \left[ (x_1^3 - x_2^2 - 3)^2 + (3x_1^2 - x_2 + 4)^2 + (x_1^4 - x_2)^2 \right]$$

cujo gradiente é dado por

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \mathbf{J}(\mathbf{x})^t \mathbf{F}(\mathbf{x}) = \begin{bmatrix} 3x_1^2(x_1^3 - x_2^2 - 3) + 6x_1(3x_1^2 - x_2 + 4) + 4x_1^3(x_1^4 - x_2) \\ -2x_2(x_1^3 - x_2^2 - 3) - (3x_1^2 - x_2 + 4) - (x_1^4 - x_2) \end{bmatrix},$$

e a matriz Hessiana de  $f$  é

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial f}{\partial x_1 \partial x_2} \\ \frac{\partial f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \mathbf{J}(\mathbf{x})^t \mathbf{J}(\mathbf{x}) + \mathbf{S}(\mathbf{x}) =$$

$$= \begin{bmatrix} 15x_1^4 - 6x_1x_2^2 - 18x_1 + 54x_1^2 - 6x_2 + 28x_1^6 - 12x_1^2x_2 + 24 & -6x_1^2x_2 - 6x_1 - 4x_1^3 \\ -6x_1^2x_2 - 6x_1 - 4x_1^3 & -2x_1^3 + 6x_2^2 + 8 \end{bmatrix}$$

As grandezas obtidas acima se fazem necessárias para a expansão truncada até o segundo termo, da série de Taylor, da função  $f$  que possibilita a solução do sistema de equações não-lineares representado por  $\mathbf{F}$ .

Aproximações quadráticas não são somente melhores que aproximações lineares, mas ganham importância quando se aproximam do ponto de solução  $\mathbf{x}^*$ . Concluí-se, observando a (6.30), que o termo de segunda ordem predomina perto do mínimo, isto por que se está buscando  $\nabla f(\mathbf{x}^*) = \mathbf{0}$  e como conseqüência  $\nabla f(\mathbf{x}^k)$  está se aproximado de zero nas proximidades do ponto de mínimo. Desta forma o gradiente da função ( $\nabla f$ ) “oscila” nas proximidades do ponto de mínimo, mudando as direções de busca à solução bruscamente. Uma conseqüência é que nas proximidades do ponto ótimo, aproximações lineares podem tornar-se instáveis.

A condição necessária de primeira ordem, para  $\mathbf{x}^+$  ser um minimizador do modelo quadrático  $m_c$  é que o seu gradiente seja nulo, ou seja,  $\nabla m_c(\mathbf{x}^+) = \mathbf{0}$ . Isto corresponde ao seguinte algoritmo para minimização irrestrita:

#### Algoritmo 6.1 (Método de Newton)

Dados  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  duas vezes diferenciável,  $\mathbf{x}^0 \in \mathbb{R}^n$  em cada iteração  $k$ , resolver,

$$\nabla^2 f(\mathbf{x}^k) \mathbf{s}_k = -\nabla f(\mathbf{x}^k)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{s}_k \tag{6.35}$$

O algoritmo 6.1 é a aplicação do método de Newton, como exposto na (6.6), considerando que neste caso constrói-se um modelo linear para o modelo quadrático  $m_c$  definido por

$$\mathbf{M}_k(\mathbf{x}^k) = \nabla m_c(\mathbf{x}^k) = \nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) \mathbf{s}_k. \quad (6.36)$$

Quando

$$\mathbf{M}_k(\mathbf{x}^k) = \nabla m_c(\mathbf{x}^k) = \mathbf{0} \quad (6.37)$$

ou de forma equivalente

$$\nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) \mathbf{s}_k = \mathbf{0} \quad (6.38)$$

fica caracterizado o passo de Newton, como exposto em (6.35).

Do ponto de vista da modelagem do problema, o algoritmo 6.1 em cada passo vai para um ponto crítico do modelo local, que pode ser de mínimo, de máximo ou ponto de sela. A minimização de  $f$  só faz sentido se  $\nabla^2 f(\mathbf{x})$  for definida positiva (ver apêndice), isto é,  $\nabla^2 f(\mathbf{x}) > 0$  para que o valor crítico do modelo seja de mínimo.

Com estas considerações é razoável tentar modificar o método de Newton para que ele manifeste preferência por minimizadores.

Os métodos para resolver (6.29) são iterativos. Para definir estes algoritmos, usam-se direções ao longo das quais é possível fazer decrescer  $f(\mathbf{x})$ , ou seja,  $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$ .

### 6.2.1 Estratégias: Busca Linear e Região de Confiança

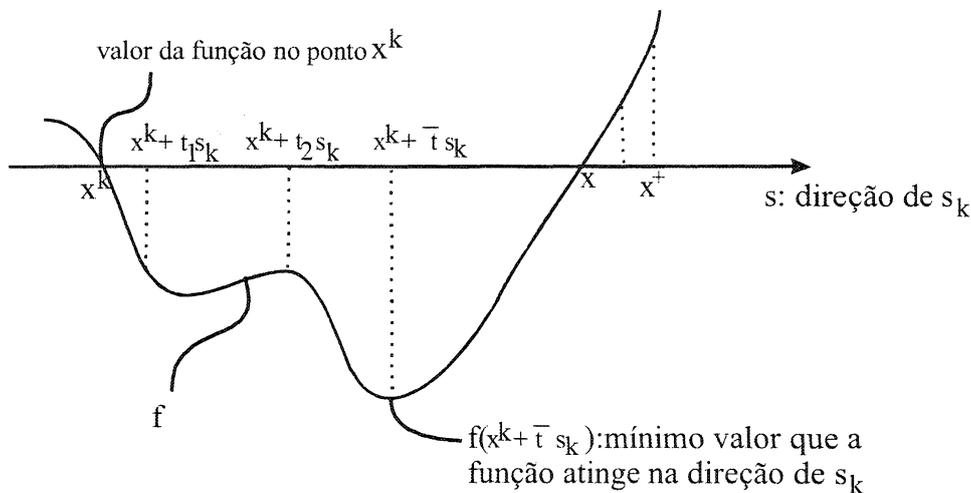
Na estratégia de busca linear, o algoritmo obtém uma direção  $\mathbf{s}_k$  e procura ao longo desta direção, a partir do ponto  $\mathbf{x}^k$ , um novo valor melhorado  $\mathbf{x}^{k+1}$  em que haja decréscimo da função objetivo. A distância que se deve mover ao longo de  $\mathbf{s}_k$ , pode

ser encontrada de forma aproximada solucionando um problema de minimização unidimensional no qual se encontra o comprimento  $t$  do passo

$$\min_{t>0} f(\mathbf{x}^k + t\mathbf{s}_k) \quad (6.39)$$

Obtendo-se  $t$  de forma exata consegue-se o menor valor que a função objetivo alcança na direção  $\mathbf{s}_k$  porém, em geral, não é necessário obtê-lo de forma exata. Muitas vezes é suficiente obter um minimizador aproximado para o problema.

FIGURA 6.4-COMPRIMENTO DO PASSO

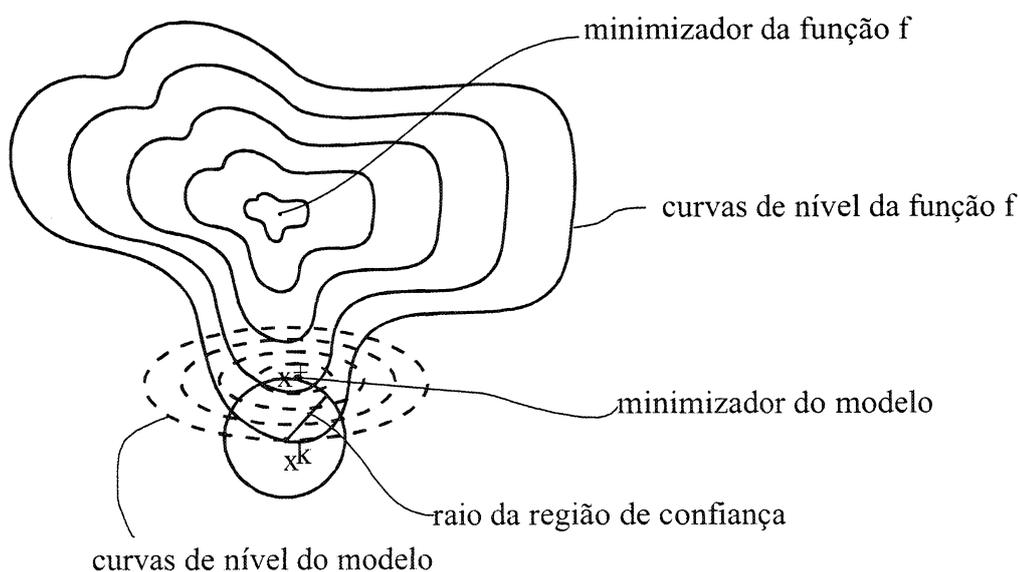


Na figura 6.4 o valor  $\bar{t}$  representa a solução exata do problema (6.39), porém observando a figura existem outros valores para  $t$ , no intervalo entre  $x^k$  e  $x$ , para os quais  $f$  alcança um valor menor que no ponto  $x^k$ , o que justifica a nomenclatura minimizador aproximado. O algoritmo de busca linear gera uma quantidade de tentativas de comprimentos de passo até encontrar um que é o minimizador aproximado de (6.39). No novo ponto  $x^{k+1} = x^k + t s_k$  uma nova direção de busca é obtida e também um novo valor para  $t$  é computado.

Uma segunda estratégia é conhecida como região de confiança. As informações obtidas sobre  $f$  são utilizadas para construir um modelo para a função, cujo comportamento em uma região próxima ao atual valor aproximado  $x^k$  é similar

ao comportamento da função  $f$  na região em torno de  $\mathbf{x}^k$  como pode ser visto na figura 6.5.

FIGURA 6.5 –REGIÃO DE CONFIANÇA



Pelo fato que o modelo quadrático pode não ser uma boa aproximação de  $f$  quando  $\mathbf{x}^+$  está distante de  $\mathbf{x}^k$ , restringe-se a procura de busca para um minimizador do modelo em alguma região em torno de  $\mathbf{x}^k$ . De forma equivalente, encontra-se um passo solucionando o subproblema, ver figura 6.5.

$$\min_{\mathbf{s}_k} m_c(\mathbf{x}^k + \mathbf{s}_k) \quad (6.40)$$

onde  $\mathbf{x}^k + \mathbf{s}_k$  esta limitado à região de confiança.

Se a solução não produz um decréscimo suficiente em  $f$ , conclui-se que a região na qual se confiou não é adequada, então esta região é reduzida e soluciona-se novamente o problema (6.40).

### 6.3 ALGORITMO DE DESCIDA COM BUSCA LINEAR

Na figura 6.4, se o método utiliza a busca linear e o passo dado é o passo de Newton, de tamanho  $\mathbf{s}_k$ , então o minimizador do modelo quadrático é representado na figura por  $\mathbf{x}^+$ , e escrito explicitamente por  $\mathbf{x}^+ = \mathbf{x}^k + \mathbf{s}_k$ . Observa-se ainda nesta figura que o valor da função no ponto  $\mathbf{x}^+$  é maior que em  $\mathbf{x}^k$ . Porém, retrocedendo o passo pode-se obter um ponto em que o valor da função  $f$  diminuí, isto é,  $f(\mathbf{x}^k) < f(\mathbf{x}^+)$ . Se o passo de Newton é utilizado e o valor da função não diminuí satisfatoriamente, em geral, se pode encontrar uma fração do tamanho do passo em que há diminuição do valor funcional. Geometricamente tomando a figura 4 como referência, a direção de  $\mathbf{s}_k$  está contida em um plano transversal que corta  $f$ .

No ponto  $\mathbf{x}^k$  se tem  $t = 0$  e em  $\mathbf{x}^+$  se tem  $t = 1$ . O valor de  $t$  que define um ponto que apresenta um decréscimo da função são os valores de  $t$  que definem pontos entre  $\mathbf{x}^k$  e  $\mathbf{x}$ . Objetiva-se um valor para  $t$  em que  $f(\mathbf{x}^k + t\mathbf{s}_k) < f(\mathbf{x}^k)$ .

#### Algoritmo 6.2 (Busca Linear)

Dados  $\mathbf{x}^0 \in \mathbb{R}^n$ ,  $\alpha \in (0,1)$ ,  $\beta > 0$ ,  $\theta \in (0,1)$ .

Para um  $\mathbf{x}^k$  inicial, a nova aproximação  $\mathbf{x}^{k+1}$  é obtida da seguinte maneira (MARTINEZ ;SANTOS, 1985; GARNÉS, 2001) :

1. Se  $\nabla f(\mathbf{x}^k) = \mathbf{0}$ , parar.

Se  $\nabla f(\mathbf{x}^k) \neq \mathbf{0}$  executar o passo 2 até o passo 5

2. escolher  $\mathbf{s}_k \in \mathbb{R}^n$  de tal forma que

$$\nabla f(\mathbf{x}^k)^t \mathbf{s}_k \leq -\theta \|\nabla f(\mathbf{x}^k)\| \|\mathbf{s}_k\| \quad (6.41)$$

$$\|\mathbf{s}_k\| \geq \beta \|\nabla f(\mathbf{x}^k)\| \quad (6.42)$$

3. fazer  $t = 1$

$$4. \text{ enquanto } f(\mathbf{x}^k + t\mathbf{s}_k) > f(\mathbf{x}^k) + \alpha t \nabla f(\mathbf{x}^k)^t \mathbf{s}_k \quad (6.43)$$

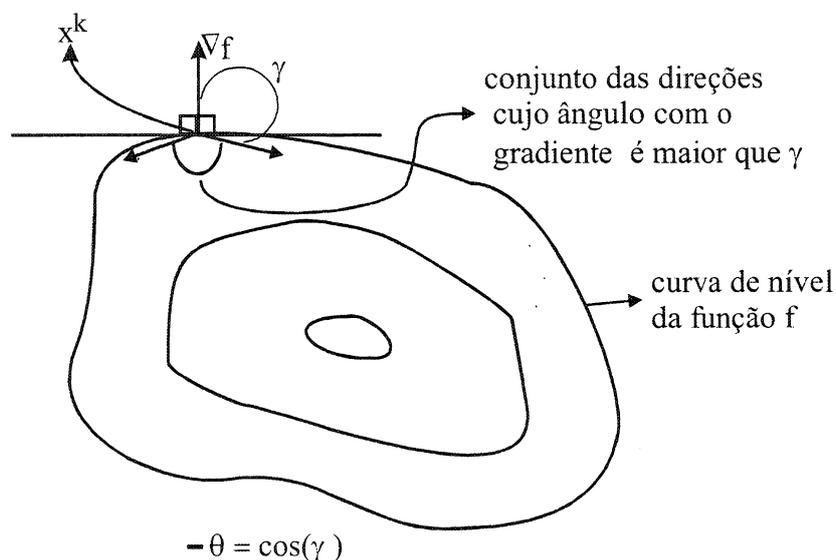
escolher um novo valor para  $t$  de forma que  $0,1\bar{t} \leq t \leq 0,9\bar{t}$ , com  $\bar{t}$  representando o valor para  $t$  fracassado.

5. fazer  $\mathbf{x}^{k+1} = \mathbf{x}^k + t\mathbf{s}_k$  e voltar para o passo 1.

Segundo MARTINEZ e SANTOS (1995) impor somente que  $f(\mathbf{x}^k + t\mathbf{s}_k) < f(\mathbf{x}^k)$ , não garante que a seqüência  $\{\mathbf{x}^k\}$  gerada pelo algoritmo convirja para um minimizador de  $f$ . Isto ocorre por que o decréscimo pode não ser suficiente, ou seja, há um decréscimo pequeno comparativamente ao comprimento do passo. A condição de Armijo (6.43) substitui o decréscimo simples, por um decréscimo mais “enérgico” de  $f$ , quando este é necessário.

Outra situação a ser evitada é que os passos dados não sejam pequenos em relação ao decréscimo de  $f$ , evitando que o método caminhe, de forma aproximada, sobre a mesma curva de nível. As direções de descida  $\mathbf{s}_k$  poderiam ser escolhidas de maneira que, à medida que o processo iterativo evolui, o ângulo entre  $\mathbf{s}_k$  e o  $\nabla f(\mathbf{x}^k)$  tendesse a  $90^\circ$ .

FIGURA 6.6-CONDIÇÃO DO ÂNGULO



O cosseno do ângulo entre  $\mathbf{s}_k$  e o  $\nabla f(\mathbf{x}^k)$ , apesar de negativo, como exige uma direção de descida, tenderia a zero. A condição do ângulo impõe que o passo seja dado em direção ao “interior” da curva de nível, como ilustrado na figura 6.6 acima. Nestas condições o ângulo entre o passo dado e o gradiente deve ser maior que  $90^\circ$  e menor que  $270^\circ$  (cosseno negativo).

Nota-se que  $\nabla^2 f(\mathbf{x}^k) \mathbf{s}_k = -\nabla f(\mathbf{x}^k)$ , então aplicando a desigualdade de norma entre vetores,

$$\|\nabla^2 f(\mathbf{x}^k)\| \|\mathbf{s}_k\| \geq \|\nabla f(\mathbf{x}^k)\| \quad (6.44)$$

e de forma equivalente

$$\|\mathbf{s}_k\| \geq \frac{1}{\|\nabla^2 f(\mathbf{x}^k)\|} \|\nabla f(\mathbf{x}^k)\| \quad (6.45)$$

comparando a desigualdade (6.45) com (6.42) se verifica que

$$\beta = \frac{1}{\|\nabla^2 f(\mathbf{x}^k)\|}. \quad (6.46)$$

A condição  $\beta$  inibe os passos muito curtos, ele impõe que o tamanho do passo seja ao menos maior que uma fração do gradiente, isto é,

$$\|\mathbf{s}_k\| \geq \beta \|\nabla f(\mathbf{x}_k)\|. \quad (6.47)$$

As direções  $\mathbf{s}_k$  que satisfazem a condição (6.47) são obtidos primeiramente pós-multiplicando a (6.47) por  $\mathbf{s}_k$

$$\|\mathbf{s}_k\| \|\mathbf{s}_k\| \geq \beta \|\nabla f(\mathbf{x}_k)\| \|\mathbf{s}_k\| \quad (6.48)$$

e posteriormente dividindo ambos os membros da (6.48) por  $\|\mathbf{s}_k\|$ , logo

$$\mathbf{s}_k \geq \beta \frac{\|\nabla f(\mathbf{x}_k)\|}{\|\mathbf{s}_k\|} \mathbf{s}_k. \quad (6.49)$$

Concluí-se que: para que a condição (6.42) seja cumprida o valor de  $\mathbf{s}_k$  deve satisfazer a desigualdade (6.49). Se  $\mathbf{s}_k$  não satisfaz a condição (6.47) é possível

corrigi-lo para que isso aconteça. Um dos valores que permitem esta correção é obtida utilizando a igualdade dada na relação (6.49), ou seja,

$$\mathbf{s}_k = \beta \frac{\|\nabla f(\mathbf{x}_k)\|}{\|\mathbf{s}_k\|} \mathbf{s}_k. \quad (6.50)$$

Com estas considerações, se o valor para  $\mathbf{s}_k$  não cumpre a condição (6.42) toma-se um novo valor para  $\mathbf{s}_k$  como dado na relação (6.50).

Quando a condição, conhecida como condição de Armijo, no passo 4 falha adota-se  $t = 0,5\bar{t}$ , onde  $\bar{t}$  é o último valor fracassado. Uma implementação mais eficiente para obter o valor de  $t$ , pode ser feita utilizando modelos quadráticos e cúbicos. A descrição detalhada desta implementação pode ser encontrada em DENNIS JR. e SCHNABEL (1983, p. 126) e NOCEDAL e WRIGHT (1999, p. 56).

O parâmetro  $\alpha$  da condição de Armijo e o parâmetro  $\theta$  são adimensionais e recomendam-se valores adequados para estes parâmetros. Usualmente  $\alpha = 10^{-4}$  ou  $10^{-1}$  e  $\theta = 10^{-6}$  ( MARTINEZ; SANTOS,1995, p. 104).

### 6.3.1 O Método de Newton Modificado com Busca Linear

O método de Newton clássico comumente aplicado, não dá preferências para minimizadores nem para maximizadores, já que a condição de otimalidade para os dois casos é a mesma, ou seja,  $\nabla f(\mathbf{x}) = 0$ . Além do mais sua aplicação não apresenta garantias quanto à solução procurada, devendo-se partir de uma vizinhança da solução (MARTINEZ; SANTOS, 1995, p. 107). Se o ponto inicial não for suficientemente “bom” não haverá garantias de convergência. Em geral, afirmações de que o ponto inicial está próximo da solução tem sentido teórico, pois, em geral, não se conhece a priori a solução e nem sequer de quão próximo pode-se estar desta solução. Portanto, partindo-se de um ponto inicial, o método pode convergir ou não.

É natural tentar modificar o método de Newton local de maneira que manifeste predileção por minimizadores e convirja independentemente do ponto inicial.

Para tanto, observa-se primeiro que, quando as direções  $\mathbf{s}_k$  são geradas como soluções de um sistema linear  $\mathbf{B}_k \mathbf{s}_k = -\nabla f(\mathbf{x}^k)$ , pré-multiplicando ambos os membros por  $\mathbf{s}_k^t$  se tem  $\mathbf{s}_k^t \mathbf{B}_k \mathbf{s}_k = -\mathbf{s}_k^t \nabla f(\mathbf{x}^k)$ . Com isso, considerando o conjunto  $\Omega = \{\mathbf{s} \in \mathbb{R}^n \mid \nabla^t f(\mathbf{x}^k) \mathbf{s} < 0\}$  os elementos de  $\Omega$  são chamados direções de descida. Portanto, direções de descida são geradas se  $\mathbf{B}_k > 0$ . Logo, a idéia é impor que as matrizes que geram direções de busca em métodos de minimização sejam definidas positivas.

O método de Newton, na solução de

$$\nabla^2 f(\mathbf{x}^k) \mathbf{s}_k = -\nabla f(\mathbf{x}^k) \quad (6.51)$$

após modificações e incorporando as condições do algoritmo 6.2 é descrito com o seguinte algoritmo (MARTINEZ; SANTOS, 1985, p. 108):

Algoritmo 6.3 (Método de Newton com Busca Linear)

Dados  $\alpha \in (0,1), \beta > 0, \theta \in (0,1)$  e  $\mathbf{x}^k \in \mathbb{R}^n$ , para  $k = 0$  escolhe-se  $\mathbf{x}^0$  e faz-se para  $k = 0,1,2,\dots$  a seguinte seqüência de passos

1. Se  $\nabla f(\mathbf{x}^k) = \mathbf{0}$ , parar

Se  $\nabla f(\mathbf{x}^k) \neq \mathbf{0}$  executar o passo 2 até o passo 8

2. tentar solucionar (5.13) usando a fatoração de Cholesky:  $\nabla^2 f(\mathbf{x}^k) = \mathbf{L}\mathbf{L}^t$

3. se a decomposição de Cholesky pôde ser realizada, obter  $\mathbf{s}_k$  resolvendo

$$\mathbf{L}\mathbf{y} = -\nabla f(\mathbf{x}^k)$$

$$\mathbf{L}^t \mathbf{s}_k = \mathbf{y}$$

4. Se no passo 2 a decomposição de Cholesky fracassou, então definir

$\mathbf{B}_k = \nabla^2 f(\mathbf{x}_k) + \mu \mathbf{I}$ ,  $\mu > 0$ . Escolher  $\mu$  de maneira que  $\mathbf{B}_k$  seja definida positiva. Com isso a decomposição de Cholesky é aceita e  $\mathbf{B}_k = \mathbf{L}\mathbf{L}^t$ . Obtém-se  $\mathbf{s}_k$  resolvendo

$$\mathbf{L}\mathbf{y} = -\nabla f(\mathbf{x}^k)$$

$$\mathbf{L}^t \mathbf{s}_k = \mathbf{y}$$

5. se  $\nabla f(\mathbf{x}^k)^t \mathbf{s}_k > -\theta \|\nabla f(\mathbf{x}^k)\| \|\mathbf{s}_k\|$  a condição do ângulo não é satisfeita, corrige-se  $\mu$  fazendo  $\mu = \text{máximo}\{2\mu, 10\}$  e repete-se o passo 4 como se tivesse havido fracasso na fatoração de Cholesky

6. Se  $\|\mathbf{s}_k\| < \beta \|\nabla f(\mathbf{x}^k)\|$  corrige-se  $\mathbf{s}_k$  com

$$\mathbf{s}_k = \beta \frac{\|\nabla f(\mathbf{x}^k)\|}{\|\mathbf{s}_k\|} \mathbf{s}_k$$

7. Enquanto

$$f(\mathbf{x}^k + t\mathbf{s}_k) > f(\mathbf{x}^k) + \alpha t \nabla f(\mathbf{x}^k)^t \mathbf{s}_k$$

$$\text{escolher } t = \frac{1}{2} \bar{t}$$

8. fazer  $\mathbf{x}^{k+1} = \mathbf{x}^k + t\mathbf{s}_k$  voltando-se ao passo 1.

No passo 4 do algoritmo acima se a matriz hessiana  $\nabla^2 f(\mathbf{x}^k)$  não é definida positiva, aumenta-se sua diagonal principal. Isto é feito escolhendo um valor para  $\mu$  suficientemente grande e utilizando a parcela  $\mu\mathbf{I}$  de forma que a positividade da matriz resultante fique garantida, ou seja, até que todos os valores próprios da matriz resultante sejam maiores que zero. Computacionalmente pode-se escolher um valor para  $\mu$  e testar a decomposição de Cholesky, se esta não é aceita escolhe um valor maior que o anterior para  $\mu$  e repete-se o processo até que a decomposição de Cholesky seja aceita.

Nestas condições, até pela positividade estar garantida, possivelmente a condição (6.41), colocada no passo 5, será satisfeita, de qualquer forma deve-se testá-la e se esta não vale, continua-se aumentando a diagonal principal da matriz hessiana, através do critério  $\mu = \text{máximo}\{2\mu, 10\}$ , até que a condição do ângulo seja satisfeita.

## 6.4 O MÉTODO DE REGIÃO DE CONFIANÇA

Método de busca linear e método de região de confiança são métodos que geram passos fazendo uso de um modelo quadrático local para uma função objetivo a qual é não linear. No entanto, estes métodos usam o modelo quadrático de forma diferente. Métodos que se baseiam em busca linear procuram usar o modelo quadrático para gerar uma direção, na qual o objetivo é obter um comprimento de passo  $t$  ao longo da direção de forma que haja um decréscimo satisfatório da função objetivo.

O método da região de confiança define uma região em torno do valor atual da iteração,  $\mathbf{x}^k$ , na qual se confia que o modelo é adequado à função objetivo, e então nesta região se obtém o passo que será o minimizador aproximado do modelo. Com efeito, este método escolhe a direção e o comprimento do passo simultaneamente. Se um passo não é aceito, reduz-se a região de confiança e encontra um novo minimizador. Em geral, a direção do passo muda quando o tamanho da região de confiança é alterado.

Para obter cada passo, em um método da região de confiança, soluciona-se o subproblema:

$$\min \mathbf{M}_k(\mathbf{s}_k) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^t \mathbf{s}_k + \frac{1}{2} \mathbf{s}_k^t \nabla^2 f(\mathbf{x}^k) \mathbf{s}_k \quad (6.52)$$

$$\text{sujeito a } \|\mathbf{s}_k\| \leq \Delta_k.$$

O teorema a seguir, cuja prova é encontrada em NOCEDAL e WRIGHT (1999) fornece uma caracterização para a solução de (6.52).

**TEOREMA 6.1** O vetor  $\mathbf{s}_k$  é uma solução do problema da região de confiança

$$\min \mathbf{M}_k(\mathbf{s}_k) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^t \mathbf{s}_k + \frac{1}{2} \mathbf{s}_k^t \nabla^2 f(\mathbf{x}^k) \mathbf{s}_k$$

$$\text{sujeito a } \|\mathbf{s}_k\| \leq \Delta_k$$

se e somente se  $\mathbf{s}_k$  é factível e existe um escalar  $\lambda \geq 0$  tal que as seguintes condições são satisfeitas:

$$\text{i) } (\nabla^2 f + \lambda \mathbf{I}) \mathbf{s}_k = -\nabla f \quad (6.53)$$

$$\text{ii) } \lambda(\Delta - \|\mathbf{s}_k\|) = 0 \quad (6.54)$$

$$\text{iii) } (\nabla^2 f + \lambda \mathbf{I}) \text{ é semi-definida positiva} \quad (6.55)$$

A condição (6.54) é conhecida como condição de complementaridade e afirma que ao menos uma das quantidades não-negativas,  $\lambda$  ou  $(\Delta - \|\mathbf{s}_k\|)$ , deve ser zero. Por isso, quando  $(\Delta - \|\mathbf{s}_k\|) > 0$  a solução de (6.52) fica estritamente no interior da região de confiança e deve-se ter  $\lambda = 0$ , portanto  $\nabla^2 f(\mathbf{x}^k) \mathbf{s}_k = -\nabla f(\mathbf{x}^k)$  que corresponde ao passo de Newton se  $\lambda = 0$ , a partir de (6.53) e (6.55),  $\nabla^2 f(\mathbf{x}^k)$  é semi-definida positiva. No caso  $(\Delta - \|\mathbf{s}_k\|) = 0$   $\lambda$  assume um valor positivo.

O Teorema 6.1 caracteriza um algoritmo para encontrar uma solução  $\mathbf{s}_k$  para (6.52). Para tanto se observa que se  $\lambda = 0$  as condições (6.53) e (6.55) são satisfeitas com  $\|\mathbf{s}_k\| \leq \Delta$ , caso contrário define-se

$$\mathbf{s}_k(\lambda) = -(\nabla^2 f + \lambda \mathbf{I})^{-1} \nabla f \quad (6.56)$$

com  $\lambda$  suficientemente grande de forma que a matriz  $\nabla^2 f + \lambda \mathbf{I}$  seja definida positiva e procura-se um valor  $\lambda > 0$  tal que

$$\|\mathbf{s}_k\| = \Delta. \quad (6.57)$$

A situação acima pode ser colocada como: deve existir um escalar  $\lambda_* > 0$  para o qual se consegue que  $\mathbf{s}_k(\lambda_*) = -(\nabla^2 f + \lambda_* \mathbf{I})^{-1} \nabla f = \Delta$ , ver figura 6.7. Este problema é um problema unidimensional, que requer a obtenção de uma raiz na variável  $\lambda$  e está descrito no algoritmo abaixo.

Algoritmo 6.4 (Região de Confiança Aproximado)

1. Escolher  $\lambda^{(0)} > 0$  e  $\Delta > 0$

Para  $\ell = 0, 1, 2, \dots$

2. obter os fatores da decomposição de Cholesky

$$\nabla^2 f + \lambda^\ell \mathbf{I} = \mathbf{R}^t \mathbf{R}$$

3. obtém-se  $\mathbf{s}_{k\ell}$ , solucionando  $(\mathbf{R}^t \mathbf{R}) \mathbf{s}_{k\ell} = \nabla f$  através da seqüência

$$\mathbf{R}^t \mathbf{y}_\ell = -\nabla f$$

$$\mathbf{R} \mathbf{s}_{k\ell} = \mathbf{y}_\ell$$

e obter  $\mathbf{q}_\ell$  através de

$$\mathbf{R}^t \mathbf{q}_\ell = \mathbf{s}_{k\ell}$$

4. Se  $\left| \|\mathbf{s}_{k\ell}\| - \Delta \right| < \varepsilon$  parar. Caso contrário ir para passo 5

5. fazer

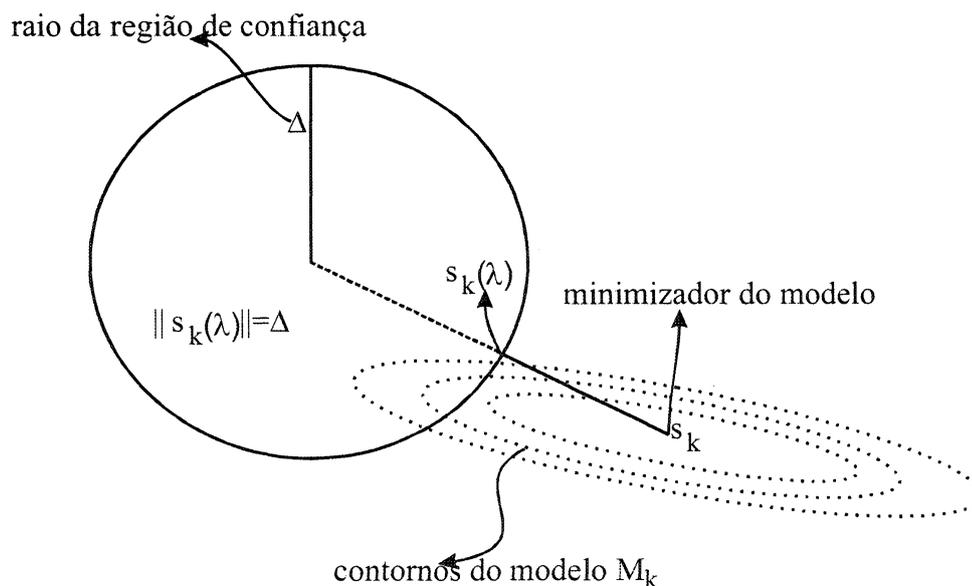
$$\lambda^{\ell+1} = \lambda^\ell + \left( \frac{\|\mathbf{s}_{k\ell}\|}{\|\mathbf{q}_\ell\|} \right)^2 \left( \frac{\|\mathbf{s}_{k\ell} - \Delta \mathbf{k}\|}{\Delta_k} \right)$$

e voltar ao passo 2.

No algoritmo 6.4 devem ser tomadas algumas salvaguardas, por exemplo, quando  $\lambda^\ell < \lambda_1$  ( $\lambda_1$  representa o menor valor próprio negativo da matriz hessiana  $\nabla^2 f$ ) nesta situação a decomposição de Cholesky não será obtida.

Outras formas de solucionar o problema (6.52) podem ser encontrados, por exemplo, em DENNIS JR e SCHNABEL (1983) e NOCEDAL e WRIGHT (1999). O valor de  $\varepsilon$  é escolhido como critério de parada para o subproblema colocado no passo 4, como sugestão  $\varepsilon = 10^{-4}$ .

FIGURA 6.7-REGIÃO DE CONFIANÇA APROXIMADO



De forma geral, descreve-se um algoritmo para solucionar o problema (6.52) considerando o algoritmo 6.4 como segue.

Se  $\nabla^2 f$  é definida positiva e  $\|s_k\| \leq \Delta$  então  $\lambda = 0$  satisfaz (6.52) e não é necessário fazer nenhuma busca, ou seja, o algoritmo 5.3 não precisa ser usado. No caso em que  $\nabla^2 f$  é positiva definida, porém  $\|s_k\| > \Delta$  existe um valor  $\lambda$  positivo no intervalo  $(0, \infty)$  e fazendo uso do algoritmo 5.3 busca-se uma solução para (6.57). No caso em que  $\nabla^2 f$  é indefinida, em geral, pode-se encontrar uma solução no intervalo  $(\lambda_1, \infty)$  se utilizando do algoritmo 6.4, onde  $\lambda_1$  representa o valor próprio mais negativo de  $\nabla^2 f$ .

Outra etapa do método de região de confiança está em escolher a região de confiança  $\Delta_k$  em cada iteração. A base para esta escolha está na conformidade entre o modelo  $M_k$  e a função objetivo  $f$  nas iterações anteriores. Dado um passo  $s_k$  define-se a razão,

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{M_k(0) - M_k(s_k)} \quad (6.58)$$

na relação (6.58) tem-se  $\mathbf{M}_k(\mathbf{s}_k) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k) \mathbf{s}_k + \frac{1}{2} \mathbf{s}_k^t \nabla^2 f(\mathbf{x}^k) \mathbf{s}_k$ .

O numerador de (6.58) é chamado de redução atual e o denominador de redução predita. Observa-se que a redução predita será sempre positiva. Desta maneira se  $\rho_k$  é negativa significa que  $f(\mathbf{x}_k + \mathbf{s}_k) > f(\mathbf{x}_k)$ , portanto o passo deve ser rejeitado.

Se  $\rho_k$  está próximo do valor 1, existe uma boa conformidade entre o modelo quadrático e a função objetivo neste passo. Então é razoável aumentar a região de confiança para a próxima iteração. Se  $\rho_k$  é positivo mas não está próximo de 1, não se altera a região de confiança a menos que  $\rho_k$  esteja próximo de zero ou negativo.

De forma geral tem-se (NOCEDAL; WRIGHT, 1999,p.68):

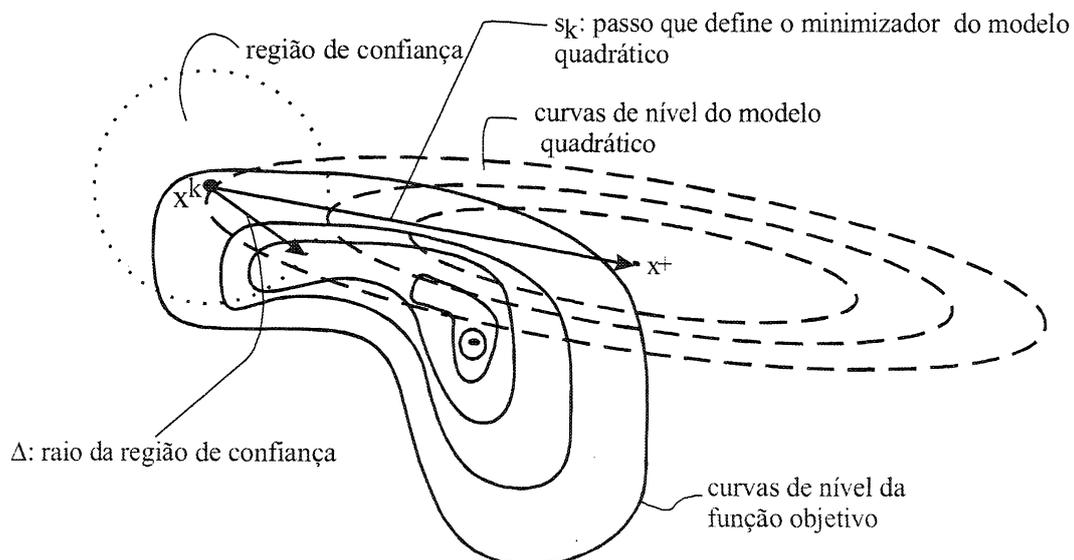
Se  $\rho_k \leq \frac{1}{4}$  o modelo é considerado “ruim”

Se  $\frac{1}{4} < \rho_k < \frac{3}{4}$  modelo é considerado razoável

Se  $\rho_k > \frac{3}{4}$  o modelo é considerado bom.

A figura 6.8 mostra o comportamento de um passo do processo iterativo quando se usa método de região de confiança e o passo de Newton.

FIGURA 6.8 – COMPARATIVO: REGIÃO DE CONFIANÇA E O PASSO DE NEWTON



Algoritmo 6.5 (Região de Confiança)

Dado  $\bar{\Delta} > 0$ ,  $\Delta_0 \in (0, \bar{\Delta})$  e  $\eta \in \left[0, \frac{1}{4}\right)$

Para  $k = 0, 1, 2, \dots$

1. Se  $\nabla f(\mathbf{x}^k) = \mathbf{0}$  parar. Senão obter  $\mathbf{s}_k$  solucionando (6.52)

2. Calcula-se  $\rho_k$  através da equação (6.58)

2.1 Se  $\rho_k < \frac{1}{4}$

Obtém-se um novo valor para a região de confiança através de

$$\Delta_{k+1} = \frac{1}{4} \|\Delta_k\|$$

e executa-se o passo 4. Caso contrário, ou seja, o passo 2.1 é contrariado, executa-se o passo 2.2.

2.2 senão, caso  $\rho_k \geq \frac{1}{4}$  ( caso em que a condição 2.1 é contrariada), testa-se as seguintes situações

$$2.2.1 \text{ se } \rho_k > \frac{3}{4} \text{ e } \|\mathbf{s}_k\| = \Delta_k$$

obtém-se um novo valor para a região de confiança calculando-se

$$\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$$

e executa-se o passo 3. Caso 2.2.1 não ocorra executa-se o passo 2.2.2.

2.2.2 senão, (caso em que a condição 2.2.1, é contrariada)

$$\Delta_{k+1} = \Delta_k$$

e executa-se o passo 5.

3. se  $\rho_k > \eta$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{s}_k$$

e executa-se o passo 5. Se a condição 3 é contrariada executa-se o passo 4

4. senão (caso a condição 3 não é verificada)

$$\mathbf{x}^{k+1} = \mathbf{x}^k$$

e vai para o passo 5.

5. volta-se em 1 e repete-se o processo até a convergência.

No Algoritmo 6.5 os passos 2.1 e 2.2 seleccionam o tamanho da região de confiança enquanto os passos 3 e 4 definem o valor para  $\mathbf{x}$  que será utilizado na próxima iteração.

No algoritmo de região de confiança, espera-se no seu desenvolvimento que o subproblema seja satisfeito com a hessiana definida positiva e que a região de confiança seja, em toda iteração, relaxada.

Em linhas gerais o algoritmo de região de confiança verifica a compatibilidade da região em torno do ponto de desenvolvimento e o tamanho do passo que se está obtendo.

## 6.5 MÉTODO QUASE NEWTON

A implementação do Método de Newton para minimizar a função (6.29), exige resolução do sistema linear dado por,

$$\left[ \mathbf{J}(\mathbf{x}^k)^t \mathbf{J}(\mathbf{x}^k) + \mathbf{S}(\mathbf{x}^k) \right] \mathbf{s}_k = -\nabla f(\mathbf{x}^k) \quad (6.59)$$

onde  $\mathbf{J}(\mathbf{x}^k)$  é dada pela relação (6.5) e  $\mathbf{S}(\mathbf{x}^k)$  pela relação (6.31) cuja solução para cada iteração é dada por,

$$\mathbf{s}_k = -\left(\nabla^2 \mathbf{F}(\mathbf{x}^k)\right)^{-1} \nabla f(\mathbf{x}^k) \quad (6.60)$$

Entretanto, esta solução às vezes exige mais de uma fatoração por iteração, a fim de que se corrija a falta de positividade da matriz Hessiana, ver passo 4 do algoritmo 6.3. Quando a dimensão do problema é “grande”, o trabalho de formar explicitamente a matriz Hessiana, pode tornar-se intolerável, e isto motiva o desenvolvimento de métodos cujo custo por iteração seja diminuído. Com isso a aplicação de algoritmos sem a necessidade das derivadas segundas se justifica.

Como o método de Newton com condições de globalização, tem boas propriedades de convergência, é esperado que os métodos “Quase Newton” tentem se parecer com ele tanto quanto o possível, mas que os custos sejam reduzidos.

A idéia é construir aproximações razoáveis da matriz Hessiana (6.33), representadas por  $\mathbf{B}_k$ , ou então da sua inversa representada por  $\mathbf{H}_k$ , ou seja,  $\mathbf{H}_k$  corresponde a  $\mathbf{B}_k^{-1}$ .

Reescrevendo a solução de (6.60) com essa consideração tem-se

$$\mathbf{s}_k = -(\mathbf{B}_k)^{-1} \nabla f(\mathbf{x}_k) \quad (6.61)$$

ou

$$\mathbf{s}_k = -\mathbf{H}_k \nabla f(\mathbf{x}^k). \quad (6.62)$$

Um algoritmo que pode ser usado nas condições acima é o Algoritmo-BFGS Globalizado, enunciado como segue (MARTINEZ; SANTOS, 1985, p. 117):

#### Algoritmo 6.6 (BFGS Globalizado)

Dados  $\alpha \in (0,1)$ ,  $\beta > 0$ ,  $\theta \in (0,1)$ ,  $\mathbf{x}^0 \in \mathbb{R}^n$ ,  $\mathbf{H}_0 = \mathbf{H}_0^t$ ,  $\mathbf{H}_0 > 0$ , (por exemplo,  $\mathbf{H}_0 = \mathbf{I}$ ) faz-se

1. se  $\nabla f(\mathbf{x}^k) = 0$  parar, caso contrário ir para o passo 2

$$2. \mathbf{s}_k = -\mathbf{H}_k \nabla f(\mathbf{x}^k),$$

3. Se  $\nabla f(\mathbf{x}^k) \mathbf{s}_k > -\theta \|\nabla f(\mathbf{x}^k)\| \|\mathbf{s}_k\|$  substitui-se  $\mathbf{s}_k$  por  $-\nabla f(\mathbf{x}^k)$  e  $\mathbf{H}_k$  por  $\mathbf{I}$ . Se

$$\|\mathbf{s}_k\| < \beta \|\nabla f(\mathbf{x}^k)\| \quad \text{deve-se substituir } \mathbf{s}_k \text{ por } \frac{\beta \|\nabla f(\mathbf{x}^k)\| \mathbf{s}_k}{\|\mathbf{s}_k\|};$$

4. escolhe-se  $t$  que satisfaça a condição de Armijo

$$f(\mathbf{x}^k + t\mathbf{s}_k) \leq f(\mathbf{x}^k) + t\alpha \nabla f(\mathbf{x}^k)^t \mathbf{s}_k \text{ e};$$

5. Define-se  $\mathbf{x}^{k+1} = \mathbf{x}^k + t\mathbf{s}_k$ ,  $\mathbf{s}_k = \mathbf{x}^{k+1} - \mathbf{x}^k$ ,  $\mathbf{y}_k = \nabla f(\mathbf{x}^{k+1}) - \nabla f(\mathbf{x}^k)$ ,

Se  $\mathbf{s}_k^t \mathbf{y}_k \leq 0$  então  $\mathbf{H}_{k+1} = \mathbf{H}_k$ , caso contrário atualiza-se a hessiana

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{(\mathbf{s}_k - \mathbf{H}_k \mathbf{y}_k) \mathbf{s}_k^t + \mathbf{s}_k (\mathbf{s}_k - \mathbf{s}_k \mathbf{H}_k \mathbf{y}_k)^t}{\mathbf{s}_k \mathbf{y}_k^t} - \frac{(\mathbf{s}_k - \mathbf{s}_k \mathbf{H}_k \mathbf{y}_k)^t \mathbf{y}_k \mathbf{s}_k \mathbf{s}_k^t}{(\mathbf{s}_k^t \mathbf{y}_k)^2}$$

6. fazer  $k = k + 1$  e voltar ao passo 1.

A fórmula BFGS tem a propriedade de produzir, em geral, matrizes definidas positivas e, portanto, com direções de descida, que frequentemente não precisarão de correção (MARTINEZ ; SANTOS, 1995, p. 116).

Uma das vantagens em se utilizar este método é que se pode iniciar como aproximação da matriz hessiana, por exemplo, a matriz identidade, e atualizá-la quando necessário.

## 6.6 MÉTODO LP PARA UM PROBLEMA DE VALOR PRÓPRIO INVERSO

Uma outra possibilidade para solucionar um problema de valor próprio inverso é a aplicação de métodos que não utilizem as derivadas primeiras e nem tão pouco das

derivadas segundas, sem perder as características de globalização requerida.

Uma forma alternativa para trabalhar com um problema de valor próprio inverso no sentido de mínimos quadrados é descrita a seguir.

Seja  $\mathbf{O}(n)$  representando todas as matrizes ortogonais no espaço  $\mathbf{R}^{n \times n}$ , e também  $\mathbf{D}^{n-m}$  representando todas as matrizes diagonais de dimensão  $(n-m) \times (n-m)$ . Para um dado conjunto

$$\mathbf{\Lambda}_m^* = \text{diag}\{\lambda_1^* \quad \dots \quad \lambda_m^*\} \quad (6.64)$$

considera-se o subconjunto.

$$\mathbf{T} = \left\{ \mathbf{Q} \text{diag}(\mathbf{\Lambda}_m^*, \mathbf{\Lambda}_{n-m}) \mathbf{Q}^t \mid \mathbf{Q} \in \mathbf{O}(n), \mathbf{\Lambda}_{n-m} \in \mathbf{D}^{n-m} \right\} \quad (6.64)$$

e também o conjunto

$$\mathbf{A} = \left\{ \mathbf{A}(\mathbf{c}) \mid \mathbf{c} \in \mathbf{R}^\ell \right\}. \quad (6.65)$$

onde  $\mathbf{A}(\mathbf{c})$  é definida como na expressão (5.1).

Observa-se, que  $\text{diag}(\mathbf{\Lambda}_m^*, \mathbf{\Lambda}_{n-m})$  é de dimensão  $n \times n$  e o conjunto  $\mathbf{T}$  representa todas as matrizes simétricas no espaço  $\mathbf{R}^{n \times n}$ , que possui os valores próprios pré-definidos  $\lambda_1^*, \dots, \lambda_m^*$  como parte do seu espectro.

Para fixar idéia, sobre os elementos do conjunto  $\mathbf{T}$ , consideram-se dois exemplos apresentados a seguir.

#### Exemplo

Dados  $n = 5$  e  $m = 3$  e o conjunto de valores próprios pré-definidos como,  $\{20 \quad 15 \quad 10\}$  o qual define a matriz  $\mathbf{\Lambda}_3^* = \text{diag}\{20 \quad 15 \quad 10\}$  fica caracterizado um elemento do conjunto  $\mathbf{T}$ , aqui representado por  $\mathbf{T}_3$ , se  $\mathbf{T}_3 \in \mathbf{T} = \left\{ \mathbf{Q} \text{diag}(\mathbf{\Lambda}_3^*, \mathbf{\Lambda}_{5-3}) \mathbf{Q}^t \mid \mathbf{Q} \in \mathbf{O}(5), \mathbf{\Lambda}_2 \in \mathbf{D}^2 \right\}$ .

Uma escolha simples para matriz ortogonal  $\mathbf{Q} \in \mathbf{O}(5)$ , é a matriz identidade de dimensão 5 e uma matriz  $\mathbf{\Lambda}_2 \in \mathbf{D}^2$  como  $\mathbf{\Lambda}_2 = \text{diag}\{4 \ 3\}$ . Considerando a matriz  $\mathbf{\Lambda}_3^* = \text{diag}\{20 \ 15 \ 10\}$  construída em função dos valores próprios pré-definidos, se obtém um elemento (matriz), representado por  $\mathbf{T}_3$ , que pertence ao conjunto  $\mathbf{T}$ ,

$$\mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{\Lambda}_3^* & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^t, \text{ que resulta em}$$

$$\mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 20 & 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^t, \text{ logo}$$

$$\mathbf{T}_3 = \begin{bmatrix} 20 & 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} \text{ a qual apresenta as características estabelecidas pelo}$$

conjunto  $\mathbf{T}$ .

Fazendo-se outras escolhas para a matriz ortogonal  $\mathbf{Q} \in \mathbf{O}(5)$  e também para matriz  $\mathbf{\Lambda}_2 \in \mathbf{D}^2$  consegue-se construir outras matrizes com  $\{20 \ 15 \ 10\}$  como parte de seu conjunto espectral.

### Exemplo

Um caso particular, e não menos importante, é o caso em que  $m = n$ . Neste caso o espectro inteiro dos elementos de  $\mathbf{T}$  são pré-definidos, ou pré-requeridos. Para exemplificar esta situação considera-se  $n = 2$  e  $\mathbf{\Lambda}_2^* = \text{diag}\{7 \ 4\}$  que define o

conjunto  $\mathbf{T}$  como  $\mathbf{T} = \left\{ \mathbf{Q} \text{diag}(\boldsymbol{\Lambda}_2^*) \mathbf{Q}^t \mid \mathbf{Q} \in \mathbf{O}(2) \right\}$ . Escolhendo a matriz ortogonal  $\mathbf{Q}$

como  $\mathbf{Q} = \begin{bmatrix} -0,995133327 & 0,098537618 \\ 0,098537618 & 0,995133327 \end{bmatrix}$  um elemento  $\mathbf{T}_2 \in \mathbf{T}$  será obtido por,

$$\mathbf{T}_2 = \begin{bmatrix} -0,995133327 & 0,098537618 \\ 0,098537618 & 0,995133327 \end{bmatrix} \times \begin{bmatrix} 7 & 0 \\ 0 & 4 \end{bmatrix} \times \begin{bmatrix} -0,995133327 & 0,098537618 \\ 0,098537618 & 0,995133327 \end{bmatrix}^t$$

$$\mathbf{T}_2 = \begin{bmatrix} 6,97087100425521 & -0,29417420260937 \\ -0,29417420260937 & 4,02912898119109 \end{bmatrix},$$

$\mathbf{T}_2$  como estabelecido acima apresenta as características pré-definidos.

Considerando a dimensão dos elementos dos conjuntos  $\mathbf{T}$  e  $\mathbf{A}$  iguais, encontrar a menor distância entre estes conjuntos é uma forma alternativa para solucionar um problema de valor próprio inverso. A formulação para este problema é dada a seguir,

Formulação III. Encontrar  $\mathbf{c} \in \mathbb{R}^\ell$ ,  $\mathbf{Q} \in \mathbf{O}(n)$  e  $\boldsymbol{\Lambda}_{n-m} \in \mathbf{D}^{n-m}$  de tal forma que a função

$$G(\mathbf{c}, \mathbf{Q}, \boldsymbol{\Lambda}_{n-m}) = \frac{1}{2} \left\| \mathbf{A}(\mathbf{c}) - \mathbf{Q} \text{diag}(\boldsymbol{\Lambda}_m^* \quad \boldsymbol{\Lambda}_{n-m}) \mathbf{Q}^t \right\|_F^2 \quad (6.66)$$

seja minimizada. Na (6.66)  $\|\bullet\|_F$  representa a norma de Frobenius (ver apêndice).

Ao problema (5.2) está associado um problema combinatório, representado por

$$\min_{1 \leq \sigma_1 \leq \dots \leq \sigma_m \leq n} \sum_{i=1}^m \left( \lambda_{\sigma_i}(\mathbf{c}) - \lambda_i^* \right)^2 \quad (6.67)$$

A representação (6.67) indica que os valores próprios pré-definidos e os valores próprios da matriz  $\mathbf{A}(\mathbf{c})$  devem estar na mesma ordem de grandeza. No caso em que  $m = n$ , isto é, quando o espectro inteiro de  $\mathbf{A}$  é pré-definido, não é necessário resolver (6.67) visto que a permutação procurada é simplesmente a permutação que arranja os valores próprios  $\lambda_i(\mathbf{c})$  na mesma ordem de  $\lambda_i^*$ .

Em linhas gerais o método LP (*lift e projection*) alterna pontos entre os conjuntos  $\mathbf{T}$  e  $\mathbf{A}$  ver figura 6.9 .

Neste método nenhuma diferenciação de valores próprios é envolvida e este converge globalmente (CHEN e CHU, 1996, p. 2422). O método LP também pode ser usado quando valores próprios múltiplos, ou seja, um conjunto de valores próprios iguais é pretendido, o que o torna uma vantagem, visto que a expressão (6.11) não está definida quando isto ocorre.

Para cada  $\mathbf{c}^k \in \mathbb{R}^\ell$ , o método LP itera os dois passos seguintes:

1. (*Lift*). Encontrar o ponto  $\mathbf{Z}^k \in \mathbf{T}$  tal que distância  $[\mathbf{A}(\mathbf{c}^k), \mathbf{Z}^k] = \text{distância}[\mathbf{A}(\mathbf{c}^k), \mathbf{T}]$ .

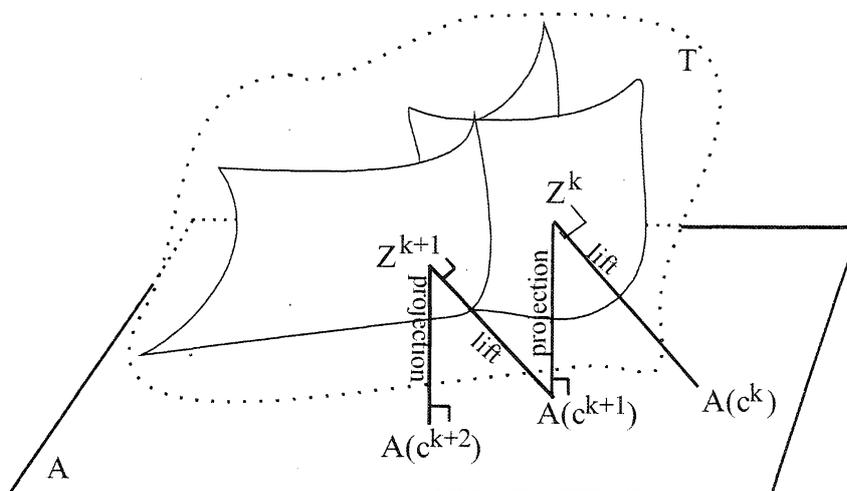
O ponto  $\mathbf{Z}^k$  é uma *lift* (levantamento) de  $\mathbf{A}(\mathbf{c}^k)$  ortogonal à  $\mathbf{T}$ .

2. (*Projection*). Encontrar o ponto  $\mathbf{c}^{k+1} \in \mathbb{R}^\ell$  tal que

distância  $[\mathbf{A}(\mathbf{c}^{k+1}), \mathbf{Z}^k] = \text{distância}[\mathbf{Z}^k, \mathbf{A}]$ . O ponto  $\mathbf{A}(\mathbf{c}^{k+1}) \in \mathbf{A}$  é chamado de *projection* (projeção) de  $\mathbf{Z}^k$  ortogonal à  $\mathbf{A}$ .

A geometria dos passos 1 e 2 é observada na figura 6.9.

FIGURA 6.9-GEOMETRIA DO MÉTODO LP



A menor distância entre  $\mathbf{A}(\mathbf{c}^k)$  e  $\mathbf{T}$ , (*Lift*), é obtido no ponto

$$\mathbf{Z}^k = \mathbf{Q}(\mathbf{c}^k) \text{diag}\left(\Lambda_m^*, \Lambda_{\bar{\sigma}^k}(\mathbf{c}^k)\right) \left[\mathbf{Q}(\mathbf{c}^k)\right]^t. \quad (6.68)$$

onde as colunas da matriz ortogonal  $\mathbf{Q}(\mathbf{c}^k)$  são os vetores próprios de  $\mathbf{A}(\mathbf{c}^k)$  arranjados de tal forma que os elementos na matriz diagonal  $\mathbf{Q}(\mathbf{c}^k) \mathbf{A}(\mathbf{c}^k) \left[\mathbf{Q}(\mathbf{c}^k)\right]^t$  estejam na mesma ordem dos elementos da matriz  $\text{diag}\left(\Lambda_m^*, \Lambda_{\bar{\sigma}^k}(\mathbf{c}^k)\right)$ . A chamada permutação  $\sigma$  reflete a reordenação dos valores e vetores próprios de  $\mathbf{A}(\mathbf{c}^k)$  e  $\bar{\sigma}^k$  representa o conjunto complementar de  $\sigma^k$ . O exemplo a seguir ilustra esta situação.

Exemplo:

Dados,  $n = 3$  e  $m = 2$ , a matriz pré-estabelecida  $\Lambda_2^* = \text{diag}(7 \ 5)$ ,

$$\mathbf{A}(\mathbf{c}) = \sum_{i=1}^2 c_i \mathbf{A}_i \text{ de dimensão } 3 \times 3, \text{ com } \mathbf{A}_1 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 6 \end{bmatrix} \text{ e } \mathbf{A}_2 = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 12 \end{bmatrix} \text{ e}$$

$\mathbf{c} = [2 \ 1]^t$ . Nas condições acima,

$$\mathbf{A}(\mathbf{c}) = \begin{bmatrix} 28 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 24 \end{bmatrix} \text{ apresenta os seguintes valores próprios } \Lambda = \text{diag}(28 \ 20 \ 24).$$

Comparando  $\Lambda = \text{diag}(28 \ 20 \ 24)$  e  $\Lambda_2^* = \text{diag}(7 \ 5)$  verifica-se que  $\Lambda_{\bar{\sigma}^k}(\mathbf{c}^k) = 28$ , pois de todas as combinações possíveis entre os elementos de  $\Lambda_2^*$  e  $\Lambda$ , as combinações  $(20, 5)$  e  $(24, 7)$  apresenta o menor valor, isto é,  $(20 - 5)^2 + (24 - 7)^2 = 514$ , seguindo (6.67).

O vetor  $\mathbf{c}^{k+1}$  que define a projeção de  $\mathbf{Z}^k \in \mathbb{R}^{n \times n}$  ortogonal à  $\mathbf{A}$  é obtido através da solução do sistema linear,

$$\sum_{i=1}^{\ell} \langle \mathbf{A}_i, \mathbf{A}_j \rangle c_i^{k+1} = \langle \mathbf{Z}^k, \mathbf{A}_j \rangle \quad j = 1, 2, \dots, \ell \quad (6.69)$$

onde  $\langle \mathbf{A}_i, \mathbf{A}_j \rangle = \text{traço}(\mathbf{A}_i^t \mathbf{A}_j)$  é denominado produto interno de Frobenius para as matrizes  $\mathbf{A}_i$  e  $\mathbf{A}_j$ .

Escrevendo (6.69) explicitamente fica,

$$\begin{bmatrix} \langle \mathbf{A}_1, \mathbf{A}_1 \rangle & \langle \mathbf{A}_2, \mathbf{A}_1 \rangle & \cdots & \langle \mathbf{A}_\ell, \mathbf{A}_1 \rangle \\ \langle \mathbf{A}_1, \mathbf{A}_2 \rangle & \langle \mathbf{A}_2, \mathbf{A}_2 \rangle & \cdots & \langle \mathbf{A}_\ell, \mathbf{A}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{A}_1, \mathbf{A}_\ell \rangle & \langle \mathbf{A}_2, \mathbf{A}_\ell \rangle & \cdots & \langle \mathbf{A}_\ell, \mathbf{A}_\ell \rangle \end{bmatrix} \begin{bmatrix} d_1^{k+1} \\ d_2^{k+1} \\ \vdots \\ d_\ell^{k+1} \end{bmatrix} = \begin{bmatrix} \langle \mathbf{Z}^k, \mathbf{A}_1 \rangle \\ \langle \mathbf{Z}^k, \mathbf{A}_2 \rangle \\ \vdots \\ \langle \mathbf{Z}^k, \mathbf{A}_\ell \rangle \end{bmatrix} \quad (6.70)$$

Observa-se que a matriz dos coeficientes no sistema linear (6.70) é independente de  $k$ , isto é, independe dos valores obtidos em cada iteração significando que esta é fixa em todo processo iterativo.

Para o caso, particular, em que o espectro inteiro de uma matriz é pretendido o algoritmo LP é proposto da seguinte forma (CHEN;CHU, 1996, p. 2424) :

#### Algoritmo 6.7 (Método LP)

Para  $k=0$ , escolhe-se um vetor  $\mathbf{c}^0 \in \mathbb{R}^n$  arbitrário e para  $k=0,1,2,\dots$  executa-se os seguintes passos

$$\mathbf{Q}(\mathbf{c}^k) \text{ e } \mathbf{\Lambda}(\mathbf{c}^k)$$

1. Forma-se  $\mathbf{A}(\mathbf{c}^k)$  e obtém-se a decomposição espectral  $\mathbf{Q}(\mathbf{c}^k)$  e  $\mathbf{\Lambda}(\mathbf{c}^k)$

onde  $\mathbf{Q}(\mathbf{c}^k)$  matriz cujas colunas são vetores próprios e  $\mathbf{\Lambda}(\mathbf{c}^k)$  matriz diagonal com valores próprios em sua diagonal principal obtidas no  $k$ -ésimo passo da iteração.

2. Forma-se a matriz  $\mathbf{Z}^k$  dada por

$$\mathbf{Z}^k = \mathbf{Q}(\mathbf{c}^k) \text{diag}(\mathbf{\Lambda}_m^*) \mathbf{Q}^t(\mathbf{c}^k)$$

onde  $\mathbf{\Lambda}_m^*$  é o vetor cujas  $m$  componentes são os valores próprios pretendidos.

3. Obtém-se  $\mathbf{c}^{k+1}$  resolvendo o sistema linear

$$\sum_{i=1}^n \langle \mathbf{A}_i, \mathbf{A}_j \rangle c_i^{k+1} = \langle \mathbf{Z}^k, \mathbf{A}_j \rangle, \quad j=1,2,\dots,n$$

4. Parar se  $\|\mathbf{c}^{k+1} - \mathbf{c}^k\| < \varepsilon$  com  $\varepsilon$  como critério de parada pretendido.

## 7 EXPERIMENTOS

Nos experimentos seguintes são apresentados alguns resultados sobre o uso dos algoritmos de otimização aqui apresentados. Os algoritmos utilizados são: método de Newton com as estratégias de busca linear e região de confiança, o método quase Newton com a aproximação da hessiana pela fórmula BFGS e o método LP (*lift and projection*). Os métodos globalizados são aqui utilizados para que se possa melhorar as garantias quanto à solução da otimização dos pesos em observações geodésicas. O método de Newton clássico já aplicado em Oliveira (2003) apesar de se apresentar como uma alternativa para a solução dos problemas de otimização de redes geodésicas quando se faz uso de um problema de valor próprio inverso, não traz garantias quanto à solução do sistema linear e como consequência poderia se concluir, falsamente, que uma determinada rede geodésica não poderia ser otimizada como se pretende.

Nas condições acima os métodos globalizados se apresentam como uma ferramenta complementar na otimização de redes geodésicas e ainda a variedade desses métodos possibilitam melhorar as chances de se encontrar os pesos, os quais permitem que a rede possa ser melhor projetada.

Os testes são simulações para redes geodésicas bidimensionais e essencialmente ilustram a aplicação dos algoritmos de otimização, ou seja, não há preocupação em verificar as variâncias das observações que os pesos obtidos representam. O valor próprio máximo pré-definido para a matriz de covariâncias, em cada exemplo, representa uma variância limite, e em consequência um desvio-padrão limite para qualquer parâmetro estimado a partir da rede.

A partir do valor limite para a variância são estabelecidos os outros valores próprios que formarão o espectro a ser objetivado na solução do problema de valor próprio inverso. Em cada exemplo aplicou-se o teste de igualdade de valores próprios a fim de verificar se a rede que se está projetando terá características de homogeneidade e isotropismo, ou seja, verifica-se se as condições (4.7) e (4.8) são válidas sob um nível de significância estabelecido.

Foram realizados três experimentos variando-se o número de valores próprios a determinar em cada um deles. Em cada experimento foram aplicados os quatro métodos apresentados, ou seja, o mesmo projeto de rede foi verificado a possível solução com cada um dos algoritmos.

Com a pluralidade de soluções (uma para cada método) utilizou-se como critério para estabelecer “a melhor solução” para um projeto de rede geodésica o critério de custo  $\sum_{j=1}^n p_j = \text{mín.}$ , ou seja, toma-se como melhor solução aquela que apresenta menor soma dos pesos. Este critério se justifica pelo fato que quanto menor o somatório dos pesos menor será o valor de cada peso em média informando que a precisão para a observação pode ser dilatada, contribuindo para uma variedade maior de instrumentos que podem ser escolhidos para o levantamento.

Todos os algoritmos foram inicializados utilizando como matriz dos pesos iniciais a matriz identidade.

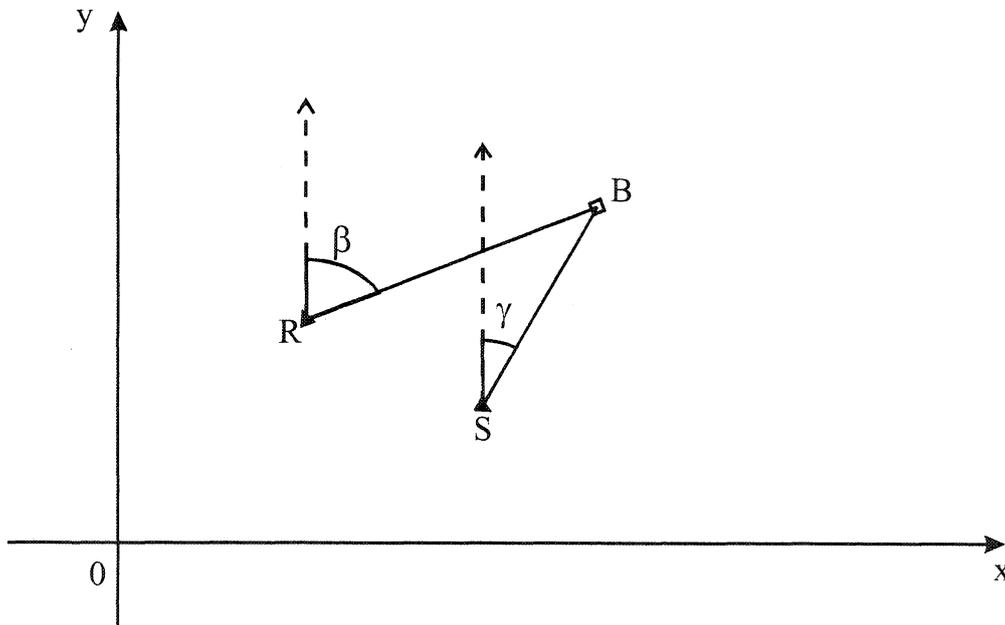
Para o método de região de confiança, após simulações com o mesmo problema, adotou-se para os três experimentos como valor “ideal”  $\Delta_0 = 0,001$ ,  $\bar{\Delta} = 10$  e como critério de parada para o método de Newton tanto com busca linear quanto com região de confiança  $\nabla f < 10^{-9}$ . Para o método quase Newton BFGS foi utilizado como aproximação inicial para a inversa da matriz hessiana  $\mathbf{H}_0$  a matriz identidade e como critério de parada também  $\nabla f < 10^{-9}$ . Para o método LP o critério de parada escolhido foi  $\|\mathbf{p}^{k+1} - \mathbf{p}^k\| < 10^{-9}$ . A seguir são apresentados os três experimentos.

## 7.1 EXPERIMENTO 1

Neste experimento estuda-se a obtenção dos pesos das observações geodésicas utilizando o método de Newton com estratégias de busca linear e região de confiança, o método de Newton com a aproximação BFGS e ainda o método LP. O problema de

valor próprio inverso será aplicado ao projeto de ajustar as coordenadas cartesianas de um ponto a partir de quatro observações como representado na figura 7.1.

FIGURA 7.1 – REPRESENTAÇÃO GEOMÉTRICA DO PLANEJAMENTO DE UM PONTO EM COORDENADAS CARTESIANAS



A representação geométrica do planeamento como apresentada na figura 7.1, estabelece que as observações planejadas serão constituídas de duas observações de distância e duas de direção.

As coordenadas do ponto  $B(x_b, y_b)$  são designadas como parâmetros e devem ser estimadas com uma precisão pré-definida estabelecida previamente por meio dos valores próprios da matriz de covariâncias dos parâmetros estimados, matriz essa obtida pela relação  $\mathbf{Q}_x = \mathbf{N}^+ = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+$ .

Para efetivação do projeto são dados:

1) As coordenadas conhecidas dos pontos R e S e as coordenadas aproximadas do ponto B apresentados na tabela 7.1.

TABELA 7.1 – COORDENADAS HORIZONTAIS I

Ponto	Coordenadas	
	x (m)	y (m)
R	200	500
S	400	300
Ponto	Coordenadas aproximadas	
B	600	582

2) O modelo funcional das 4 observações divididas em:

Duas observações de distância:

$$d_1 = \sqrt{(x_b - 200)^2 + (y_b - 500)^2}$$

$$d_2 = \sqrt{(x_b - 400)^2 + (y_b - 300)^2}$$

duas observações de direção:

$$\beta = \arctan\left(\frac{x_b - 200}{y_b - 500}\right)$$

$$\gamma = \arctan\left(\frac{x_b - 400}{y_b - 300}\right)$$

3) A matriz de planejamento **A** é obtida das derivadas parciais das equações de observação, avaliadas nas coordenadas aproximadas do ponto B. Para a rede representada pela figura 14 é

$$\mathbf{A} = \begin{bmatrix} 0,97963 & 0,20082 \\ 0,57850 & 0,81568 \\ 0,00049 & -0,00240 \\ 0,00236 & -0,00167 \end{bmatrix}$$

Apresentando quatro linhas (número de observações) e duas colunas (número de parâmetros).

4) Valores próprios pré-estabelecidos para a matriz de covariâncias são:

$$\mu_1 = \frac{1}{15000} \text{m}^2 \text{ e } \mu_2 = \frac{1}{20000} \text{m}^2.$$

Desta forma o valor limite para a variância de qualquer parâmetro estimado nesta rede é de  $\frac{1}{15000} \text{m}^2 \cong 0,0000666 \text{m}^2$ , que é o valor próprio máximo estabelecido para a matriz de covariâncias.

#### 5) Aplicação do teste de igualdade dos valores próprios.

Para este caso, que é o caso bivariado, aplicou-se tanto a estatística dada em (4.5) como a dada em (4.6) e ambas com um nível de significância para o teste de 5%. Os dois testes forneceram a mesma conclusão, ou seja, indicaram que a hipótese da rede projetada ser homogênea e isotrópica não deve ser rejeitada, isto com um nível de confiança de 95% para o teste.

6) Os valores próprios utilizados no processo de otimização estabelecido pela equação (4.7) são

$$\lambda_1^* = 20000 \text{ e } \lambda_2^* = 15000.$$

A matriz  $\mathbf{P}$  é uma matriz diagonal cujos elementos na diagonal principal são tomados como incógnitas.

### PROBLEMA 1

O problema é enunciado como: quais pesos ( ou qual matriz de pesos) deve ser atribuído às observações para que a matriz dos coeficientes das equações normais (matriz  $\mathbf{N}$ ) apresente os valores próprios  $\lambda_1 = 20000$  e  $\lambda_2 = 15000$ ?

A matriz  $\mathbf{N}$  é escrita como,

$$\mathbf{N} = \sum_{j=1}^4 \left( p_j \mathbf{a}_j \mathbf{a}_j^t \right)$$

onde  $p_j = p_{jj}$  é cada elemento da diagonal principal da matriz  $\mathbf{P}$  e  $\mathbf{a}_j$  é a  $j$ -ésima linha da matriz  $\mathbf{A}$ .

Para solucionar o problema 1 fez-se uso dos métodos de Newton com estratégias de busca linear e região de confiança, método quase Newton BFGS e o Método LP.

Os resultados estão colocados na tabela 7.2:

TABELA 7.2- RESULTADO DA OTIMIZAÇÃO: BUSCA LINEAR, REGIÃO DE CONFIANÇA, BFGS E LP.

<b>Método</b>	<b>Pesos</b>	<b>Número de iterações</b>
LP	8201,23996 12720,97180 796778022,61755 1111425479,72264	2
Newton com busca linear	15784,03949 5138,17227 2061012691,53421 205098948,64417	2
Newton com região de confiança	15784,03949 5138,17227 2061012691,53421 205098948,64417	2
BFGS	divergiu	-

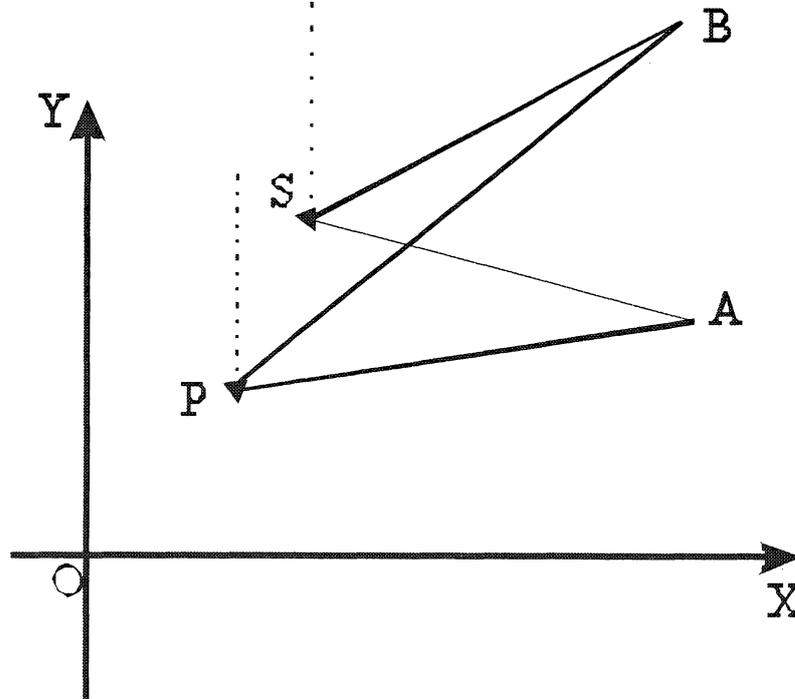
Para esta exemplificação os métodos: LP, Método de Newton com busca linear e região de confiança alcançaram a solução com o mesmo número de iterações. Apenas o método BFGS não alcançou a solução, divergindo. No caso de uma otimização de pesos qualquer solução pode ser usada, porém pode-se tomar como critério de escolha a solução que apresentou menores pesos, pois elas requerem menos precisão para sua efetivação e em consequência instrumentos menos precisos que pode ser um fator de minimização de custo.

## 7.2 EXPERIMENTO 2

Neste experimento, a otimização dos pesos será aplicada em uma rede horizontal com o propósito de densificá-la a partir de 2 pontos conhecidos. O projeto

prevê dois pontos novos a serem determinados, com precisão previamente estabelecida.

FIGURA 7.2 – PROJETO DE REDES GEODÉSICAS: GEOMETRIA PARA OBTENÇÃO DE DOIS PONTOS EM COORDENADA HORIZONTALIS



Considera-se o projeto, cuja configuração geométrica é dada na figura 7.2, com os pontos S e P sendo fixos (considerados isentos de erros) e os pontos A e B a serem determinados, por meio de quatro observações de distância e quatro observações de azimute (ângulo).

Para execução deste projeto são dados:

- 1) As coordenadas conhecidas dos pontos P e S e as coordenadas aproximadas dos pontos A e B colocadas na tabela 7.3.

TABELA 7.3 - COORDENADAS HORIZONTAIS II

Ponto	Coordenadas	
	x (m)	y (m)
P	200	150
S	300	400
Ponto	Coordenadas aproximadas	
A	550	275
B	460	580

2) O modelo funcional das 8 observações divididas em:

Quatro observações de distância:

$$d_{PA} = \sqrt{(x_a - 200)^2 + (y_a - 150)^2}$$

$$d_{SA} = \sqrt{(x_a - 300)^2 + (y_a - 400)^2}$$

$$d_{SB} = \sqrt{(x_b - 300)^2 + (y_b - 400)^2}$$

$$d_{PB} = \sqrt{(x_b - 200)^2 + (y_b - 150)^2}$$

e quatro observações de azimute:

$$\alpha_{PA} = \arctan\left(\frac{x_a - 200}{y_a - 150}\right)$$

$$\alpha_{PB} = \arctan\left(\frac{x_b - 200}{y_b - 150}\right)$$

$$\alpha_{SA} = \arctan\left(\frac{x_a - 300}{y_a - 400}\right)$$

$$\alpha_{SB} = \arctan\left(\frac{x_b - 300}{y_b - 400}\right)$$

3) A matriz  $\mathbf{A}$  é obtida avaliando-se as derivadas parciais de cada modelo funcional em relação aos parâmetros em um ponto inicial, representado por  $\mathbf{X}_0$ , que neste caso são os valores aproximados da tabela 3. Para este exemplo a matriz planejamento  $\mathbf{A}$  tem 8 linhas ( número de observações) e 4 colunas ( número de parâmetros). E escrita explicitamente fica:

$$\mathbf{A} = \begin{bmatrix} 0,9701 & 0,2425 & 0 & 0 \\ 0,8944 & -0,4472 & 0 & 0 \\ 0 & 0 & 0,7682 & 0,6402 \\ 0 & 0 & 0,6247 & 0,7809 \\ 0,0006 & -0,0024 & 0 & 0 \\ 0 & 0 & 0,0012 & -0,0010 \\ -0,0013 & -0,0027 & 0 & 0 \\ 0 & 0 & 0,0016 & -0,0020 \end{bmatrix}$$

4) Valores próprios pré-estabelecidos para a matriz de covariâncias são:

$$\mu_1 = \frac{1}{10000} \text{ m}^2 \quad \mu_2 = \frac{1}{12000} \text{ m}^2 \quad \mu_3 = \frac{1}{15000} \text{ m}^2 \quad \text{e} \quad \mu_4 = \frac{1}{20000} \text{ m}^2.$$

Desta forma o valor limite para a variância de qualquer parâmetro estimado nesta rede é de  $\frac{1}{10000} \text{ m}^2 = 0,0001 \text{ m}^2$ , que é o valor próprio máximo estabelecido para a matriz de covariâncias.

5) Aplicação do teste de igualdade dos valores próprios.

Para este caso aplicou-se a estatística dada em (4.5) com um nível de significância para o teste de 5%. O resultado indicou que a hipótese da rede projetada ser homogênea e isotrópica não é rejeitada com um nível de confiança de 95%.

6) Os valores próprios utilizados no processo de otimização estabelecido pela equação (4.7) são

$$\lambda_1^* = 20000 \quad \lambda_2^* = 15000 \quad \lambda_3^* = 12000 \quad \lambda_4^* = 10000$$

A matriz  $\mathbf{P}$  é uma matriz diagonal cujos elementos na diagonal principal são tomados como incógnitas.

#### PROBLEMA 2

O problema é enunciado como: quais pesos ( ou qual matriz de pesos) deve ser atribuído às observações para que a matriz dos coeficientes das equações normais (matriz  $\mathbf{N}$ ) apresente os valores próprios  $\lambda_1 = 20000$  e  $\lambda_2 = 15000$   $\lambda_3 = 12000$   $\lambda_4 = 10000$ ?

A matriz  $\mathbf{N}$  é escrita como,

$$\mathbf{N} = \sum_{j=1}^8 \left( p_j \mathbf{a}_j \mathbf{a}_j^t \right)$$

como anteriormente  $p_j = p_{jj}$  é cada elemento da diagonal principal da matriz  $\mathbf{P}$  e  $\mathbf{a}_j$  é a  $j$ -ésima linha da matriz  $\mathbf{A}$ .

Para solucionar o problema 2 novamente os 4 métodos propostos foram utilizados e o peso obtido por cada método para cada observação planejada é apresentado no quadro 7.1

QUADRO 7.1- RESULTADO DO PROCESSO DE OTIMIZAÇÃO DOS PESOS

<b>Método de Newton: busca linear</b>	<b>Método de Newton: Região de confiança</b>	<b>Método LP</b>
14693,04650	-11923,12420	5485,45146
782,57017	27398,74087	9990,16493
16368,54908	2022,65305	13797,17737
3734,51204	18080,40808	6305,88376
2161963064,77249	-2362786051,55650	59667151,943616
-561105138,60352	5320711754,67274	493153114,98274
-134217726,99742	2860101541,16148	901637086,29927
1717986919,40156	-469762046,99997	1325854276,30006

O método BFGS foi aplicado, mas não alcançou solução, divergindo.

Do ponto de vista da solução do processo de otimização o método de Newton com busca linear indica que duas observações de direção devem ser eliminadas, pois

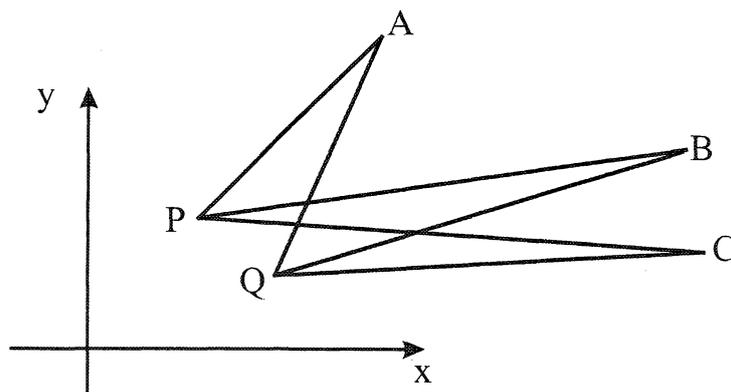
não cooperam com a precisão final da rede. A solução do método de região de confiança indica que três observações devem ser eliminadas, uma de distância e duas de direção. O método LP indica que todas as observações cooperam com a precisão e, portanto nenhuma delas deve ser descartada.

### 7.3 EXPERIMENTO 3

No exemplo a seguir formula-se um problema de valor próprio inverso aplicado à Lei de Propagação de Covariâncias utilizada após um ajustamento por mínimos quadrados na sua forma paramétrica.

Na figura 7.3 está representada uma rede geodésica bidimensional onde as coordenadas dos pontos  $P(x_p, y_p)$  e  $Q(x_q, y_q)$  são conhecidas e consideradas isentas de erros.

FIGURA 7.3- REDE GEODÉSICA BIDIMENSIONAL



As coordenadas dos pontos  $A(x_a, y_a)$ ,  $B(x_b, y_b)$  e  $C(x_c, y_c)$ , designadas de parâmetros, devem ser estimadas com uma precisão pré-definida dada pelos valores próprios da matriz de covariâncias dos parâmetros estimados, obtida pela relação  $\mathbf{Q} = \mathbf{N}^+ = (\mathbf{A}^t \mathbf{P} \mathbf{A})^+$ .

Dados:

1) as coordenadas dos pontos fixos e também as coordenadas aproximadas aos parâmetros estão apresentadas na tabela 7.4

TABELA 7.4- COORDENADAS HORIZONTAIS DE UMA REDE GEODÉSICA BIDIMENSIONAL

pontos	Coordenadas dos Pontos Fixos	
	x (m)	y (m)
P	2,1	2,5
Q	3,5	1,3
pontos	Valores Aproximados	
	x (m)	y (m)
A	5,0	6,0
B	8,4	4,8
C	8,9	2,4

2) O modelo funcional das 12 observações, divididas em: seis observações de distância

$$d_{PA} = \sqrt{(2,1 - x_a)^2 + (2,5 - y_a)^2}$$

$$d_{PB} = \sqrt{(2,1 - x_b)^2 + (2,5 - y_b)^2}$$

$$d_{PC} = \sqrt{(2,1 - x_c)^2 + (2,5 - y_c)^2}$$

$$d_{QA} = \sqrt{(3,5 - x_a)^2 + (1,3 - y_a)^2}$$

$$d_{QB} = \sqrt{(3,5 - x_b)^2 + (1,3 - y_b)^2}$$

$$d_{QC} = \sqrt{(3,5 - x_c)^2 + (1,3 - y_c)^2}$$

e 6 observações de azimute.

$$\alpha_{PA} = a \tan \left( \frac{x_a - 2,1}{y_a - 2,5} \right)$$

$$\alpha_{PB} = a \tan \left( \frac{x_b - 2,1}{y_b - 2,5} \right)$$

$$\alpha_{PC} = a \tan \left( \frac{x_c - 2,1}{y_c - 2,5} \right)$$

$$\alpha_{QA} = a \tan \left( \frac{x_a - 3,5}{y_a - 1,3} \right)$$

$$\alpha_{QB} = a \tan \left( \frac{x_b - 3,5}{y_b - 1,3} \right)$$

$$\alpha_{QC} = a \tan \left( \frac{x_c - 3,5}{y_c - 1,3} \right)$$

3) A matriz **A** é obtida avaliando-se as derivadas parciais de cada modelo funcional em relação aos parâmetros em um ponto inicial, aqui representado por  $\mathbf{X}_0$ , que neste caso é dado pelos valores aproximados da tabela 1. Para o exemplo a matriz **A** apresenta 12 linhas (número de observações) e 6 colunas (número de parâmetros). A matriz **A** escrita explicitamente fica

$$\mathbf{A} = \begin{bmatrix} 0,6380 & 0,7700 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,9394 & 0,3429 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,9999 & -0,0147 \\ 0,3040 & 0,9527 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,8137 & 0,5812 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,9799 & 0,1996 \\ 0,1694 & -0,1404 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,0511 & -0,1401 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0,0022 & -0,1470 \\ 0,1931 & -0,0616 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,0965 & -0,1351 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,0362 & -0,1778 \end{bmatrix}$$

4) Valores próprios pré-estabelecidos para a matriz de covariâncias são:

$$\mu_1 = \frac{1}{10000} \text{ m}^2 \quad \mu_2 = \frac{1}{20000} \text{ m}^2 \quad \mu_3 = \frac{1}{30000} \text{ m}^2 \quad \mu_4 = \frac{1}{40000} \text{ m}^2$$

$$\mu_5 = \frac{1}{50000} \text{ m}^2 \quad \text{e} \quad \mu_6 = \frac{1}{60000} \text{ m}^2 .$$

Desta forma o valor limite para a variância de qualquer parâmetro estimado nesta rede é de  $\frac{1}{10000} \text{m}^2 = 0,0001 \text{m}^2$ , que é o valor próprio máximo estabelecido para a matriz de covariâncias.

5) Aplicação do teste de igualdade dos valores próprios.

Para este caso aplicou-se a estatística dada em (4.5) com um nível de significância para o teste de 5%. O resultado indicou que a hipótese da rede projetada ser homogênea e isotrópica não é rejeitada com um nível de confiança de 95%.

6) Os valores próprios utilizados no processo de otimização estabelecido pela equação (4.7) são

$$\lambda_1^* = 60000 \quad \lambda_2^* = 50000 \quad \lambda_3^* = 40000 \quad \lambda_4^* = 30000 \quad \lambda_5^* = 20000 \quad \lambda_6^* = 10000.$$

A matriz  $\mathbf{P}$  é uma matriz diagonal cujos elementos na diagonal principal são tomados como incógnitas.

### PROBLEMA 3

O problema é enunciado como: quais pesos ( ou qual matriz de pesos) deve ser atribuído às observações para que a matriz dos coeficientes das equações normais (matriz  $\mathbf{N}$ ) apresente os valores próprios  $\lambda_1 = 60000$  e  $\lambda_2 = 50000$   $\lambda_3 = 40000$   $\lambda_4 = 30000$   $\lambda_5 = 20000$   $\lambda_6 = 10000$ ?

A matriz  $\mathbf{N}$  é escrita como,

$$\mathbf{N} = \sum_{j=1}^{12} \left( p_j \mathbf{a}_j \mathbf{a}_j^t \right)$$

como anteriormente  $p_j = p_{jj}$  é cada elemento da diagonal principal da matriz  $\mathbf{P}$  e  $\mathbf{a}_j$  é a  $j$ -ésima linha da matriz  $\mathbf{A}$ .

Com estas considerações formula-se o problema de valor próprio inverso como dado na (5.1) aplicado na obtenção de pesos otimizados para uma rede geodésica como:

Conhecida a matriz  $\mathbf{A}$  proveniente de um ajustamento por mínimos quadrados, obter os pesos para as observações geodésicas para que se tenha como valores próprios da matriz  $\mathbf{N}$  os valores:

$$\lambda_1^* = 60000 \quad \lambda_2^* = 50000 \quad \lambda_3^* = 40000 \quad \lambda_4^* = 30000 \quad \lambda_5^* = 20000 \quad \lambda_6^* = 10000.$$

O mesmo planejamento foi solucionado com o Método quase Newton, BFGS e o método de Newton com região de confiança e busca linear e ainda o método LP os que apresentaram os seguintes pesos para o planejamento mostrado no quadro 7.2:

QUADRO 7.2 – PESOS OBTIDOS PELOS MÉTODOS: LP, BFGS E REGIÃO DE CONFIANÇA

<b>Método LP</b>	<b>Método quase Newton BFGS</b>	<b>Região de confiança</b>
22575,258548178	32827,537120012	17523,162623486
23263,903609251	24544,866942047	25014,412134193
28576,863675617	47126,907435123	26329,257480257
17815,273070396	28344,031030285	22867,374156731
27300,838442210	26208,087336893	25501,176185199
32015,959280592	-9889,596983932	34263,586075533
355463,046649752	346089,649965465	250848,841630837
344388,938341000	163852,474331488	627988,884741480
131133,227623487	316318,214151395	33495,465311296
301915,633905575	293948,206834533	425163,802083439
427098,111189884	203210,489658378	200259,820965537
199587,510108045	483592,954525169	263700,589178145

O método de Newton com busca linear apresentou um número excessivamente grande de iterações para alcançar uma solução, por isso aqui sua solução não foi apresentada.

Os métodos LP, BFGS e região de confiança alcançaram a solução. Salienta-se aqui que o método LP utilizou apenas 2 iterações para obter a solução enquanto o Método BFGS utilizou 6551 iterações e o método de região de confiança 550 iterações.

## 8 CONCLUSÃO

A otimização dos pesos das observações geodésicas é realizável tanto quando se usa o método baseado em matriz critério quanto o método baseado em valores próprios. Ambos os métodos procuram os pesos baseados essencialmente em uma precisão pré-definida para a rede, precisão esta estabelecida sobre alguma propriedade ótima para a matriz de covariâncias. O diferencial entre as estratégias reside na forma em que é postulada tal precisão. Das duas estratégias a mais simples, ao menos do ponto de vista do estabelecimento do critério ótimo para a rede, é a estratégia baseada no valor próprio. Neste caso é necessário definir somente um valor para o critério de precisão, ou seja, a precisão limite para qualquer parâmetro estimado na rede o qual coincide com o valor próprio máximo da matriz de covariâncias dos parâmetros estimados, os demais valores serão necessariamente menores que este valor máximo exigindo-se somente que o conjunto todo seja homogêneo e isotrópico, condição esta que pode ser verificada baseando-se nos critérios de otimalidade de uma rede geodésica e pela aplicação do teste de igualdade de valores próprios.

Em contrapartida quando se usa matriz critério deve-se obter  $u^2$  (  $u$  sendo o número de parâmetros) elementos para a matriz critério de covariância. Elementos estes que podem ser obtidos por modelos para a matriz critério ou de forma arbitrária baseado na experiência do projetista.

Do ponto de vista dos métodos analíticos para a otimização dos pesos, ambas as estratégias apresentam resultados satisfatórios nos processos de otimização.

Especificamente quando é utilizado um problema de valor próprio inverso para a obtenção dos pesos tem-se um problema não linear e a utilização dos algoritmos de otimização se faz necessário.

De forma geral os algoritmos, método de Newton com busca linear ou região de confiança, método quase-Newton BFGS e o método LP podem ser utilizados na busca pela solução do problema, porém o método menos confiável a conduzir para a solução é o método BFGS, que em dois dos três experimentos (1 e 2) não alcançou solução, divergindo.

O método de Newton com busca linear obteve resposta em todos os casos sendo desconsiderado no experimento 3 pelo número excessivo de iterações que se fez necessário para levar à uma solução ( neste caso se utilizou o critério de parada 10000 iterações).

O método de Newton com região de confiança alcançou solução em todos os casos aplicados.

Nas aplicações o método que obteve os melhores resultados foi o método LP, em todos os casos obteve solução com apenas duas iterações. Provavelmente deve-se ao fato deste método ser exclusivo para um problema de valor próprio inverso e de não necessitar derivadas primeiras e nem de derivadas segundas forçando o método a ser “estável” em todos os passos da iteração. Fato que pode não ocorrer com os métodos que se utiliza, de derivadas segundas haja vista a equação (6.11) não estar definida para valores próprios iguais ou aproximadamente iguais, fato que pode ocorrer em algum passo do processo iterativo.

A situação acima justifica a aplicação do teste de igualdade dos valores próprios antes do processo de otimização, ou seja, escolhem-se valores próprios iguais entre si sob um nível de significância possibilitando que mesmo assim a equação (6.11) esteja definida o que não ocorreria se os valores próprios fossem escolhidos matematicamente iguais entre si.

Desta forma baseando-se nos experimentos pode-se estabelecer que, em ordem de eficiência, os melhores métodos para solução irrestrita de um problema de valor próprio inverso aplicado ao problema de otimização dos pesos das observações geodésicas são: o método LP, o método de Newton com região de confiança, o método de Newton com busca linear e o método quase-Newton BFGS.

Considerando o critério de custo para uma rede geodésica ainda pode-se concluir que no experimento 1 das soluções apresentadas pelos métodos LP, Newton por busca linear e Newton com região de confiança o que minimiza o custo para a precisão das observações da rede é a solução obtida pelo método LP. Para este método

$\sum_{j=1}^4 p_j = 1908224424,55195$ . Enquanto que para os outros dois métodos

$$\sum_{j=1}^4 p_j = 2266132562,317877.$$

No experimento 2 para o qual se alcançou solução com o método de Newton com busca linear e região de confiança e ainda o método LP obteve-se para o critério de custo, respectivamente, (calculados somente para os pesos positivos visto que observações com pesos negativos devem ser retiradas do projeto):

$$\sum_{j=1}^8 p_j = 387998562,851840 \quad , \quad \sum_{j=1}^8 p_j = 8180860797,63622 \quad e$$

$$\sum_{j=1}^8 p_j = 2780347208,203206.$$

O método que apresentou melhor eficiência quanto aos custos de precisão das observações para a rede é o método LP seguido do método por região de confiança sendo o método com maior custo o método com busca linear.

No experimento 3 os métodos que alcançaram solução foram: o método LP, o método BFGS e o método de região de confiança, cujas somatório dos pesos

$$\text{respectivamente são, } \sum_{j=1}^{12} p_j = 1569956,372566133 \quad \sum_{j=1}^{12} p_j = 1966063,419330788 \quad e$$

$$\sum_{j=1}^{12} p_j = 1911134,564443987, \text{ sendo o que apresentou um menor custo do ponto de}$$

vista da precisão da observação foi o método BFGS.

## 9 REFERÊNCIAS

- AMIRI-SIMKOOEEI, ALIREZA. (2004). A new method for second order design of geodetic networks: aiming at high reliability. **Survey Review**, v. 37, n. 293. pp.552-560.
- BAARDA, W. (1962). A generalization of the concept Strength of Figure. **Publ. of the Computing Center of the Geodetic Institute**. Delft.
- BIEGLER-KÖNIG, F. W. (1981a). A newton iteration process for inverse eigenvalue problem. **Numerische Mathematik**. V. 37 p. 349-354.
- BIEGLER-KÖNIG, F. W. (1981b). Sufficient conditions for the solubility of inverse eigenvalue problem. **Linear Algebra and Applications**. n 40 pp. 89-100.
- BRONSON, R. (1993). **Matrizes**. Lisboa: McGaw-Hill.
- CHEN, X; CHU, M. T.(1996). On least square solution of inverse eigenvalue problem. **SIAM Journal Numerical Analyses**. v. 33 n. 6. pp. 2417-2430.
- DENNIS Jr., J. E.; SCHNABEL, R. B. (1983). **Numerical methods for unconstrained optimization and nonlinear equations**. Englewood Cliffs: Prentice-Hall.
- DEREN, L.; YONGQIAN, Z. (1991). Optimization and design of geodetic network in consideration of accuracy and reliability. **Allgemeine Vermessung-Nachrichten**, Karlsruhe, v. 91, n. 8, p. 27-33.
- DOWNING, A.C.; HOUSEHOLDER, A.S.(1956). Some inverse characteristic value problem. **J. ACM**. v 3 pp 203-207.
- DUPRAZ, H. ; NIEMEIER, W. (1981). **Beurteilungskriterien für Geodätischer Netze**. In: Beiträge zum II. Internationalen Symposium über Deformationsmessungen mit Geodätischen Methoden. Bonn 25-28. September 1978. Herausgegeben von Ludger Hallermann, Bonn. Stuttgart: Wittwer, S., 386-400.
- FRIEDLAND, S.; NOCEDAL, J.; OVERTON, M. L. (1987). The formulation and analysis of numerical methods for inverse eigenvalue problems. **SIAM J. Numer. Anal.**, v. 24, n. 3, p. 634-667.
- FRIEDLAND, S. (1977). Inverse eigenvalue problem. **Linear Algebra and Applications**. n 17 pp 15-51.
- FRIEDLANDER, A. (1994). **Elementos de programação não-linear**. Campinas: Editora da UNICAMP.

GARNÉS, S. J. A. (2001). **Resolução das ambigüidades GPS para linhas de base curta: análise dos algoritmos de otimização**. Curitiba, 2001. Tese (Doutorado em Ciências Geodésicas) – Departamento de Geociências, Universidade Federal do Paraná

GEMAEL, C.(1994). **Introdução ao ajustamento de observações: aplicações geodésicas**. Curitiba: Universidade Federal do Paraná.

GRAFAREND, E. W. (1972). Genauigkeitsmasse geodatischer Netze. **Deutsche Geodätische Kommission**. Publ. A-73, Munchen.

GRAFAREND, E. W. ; SCHAFRIN, B. (1979). Kriterion-matrizen I-zweidimensional homogene und isotropoe geodatische netze. **Zeitschrift Für Vermessungswesen**.v. 4 pp. 133-149.

GRAFAREND, E. W. ; SANZO F. (1985). **Optimization and design of geodetic networks**. Spring Verlag, Berlin-Heidelberg-New York-Tokio

HELMERT, F.R. (1868). Studien über rationelle Vermessungen im Gebiet der höheren Geodäsie. **Z. Math. Phys. Schömilch**. v. 13 pp 73-120 e 163-168.

HENDERSON, H. V.; SEARLE,S.R. (1979). Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics. **Canadian Journal of Statistics**, v. 7, n. 1, p. 65-81.

JACKSON, J. E. (1991). **A user's guide to principal componentes**. New York. J. Wiley.

JÄGER, R. (1988). Analyse und Optimierung geodätischer Netze nach spektralen Kriterien und mechanische Analogien. **Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften**, München, série C, n. 342.

JUNG, I. (1924). Über die günstigste Gewisheitsverteilung in Basisnetzen. **Akadem. Abh.** Uppsala 1924

KALTENBACH, H. (1992). Optimierung geodätischer Netze mit spektralen Zielfunktionen. **Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften**, München, série C, n. 393.

KUANG, S. (1996). **Geodetic network analysis and optimal design: concepts and applications**. Chelsea: Ann Arbor Press.

KUBLANOVSKAJA, W.N.(1970). **On an approach to the solution of the inverse eigenvalue problem**. Zapinski naucnych Seminarov Leningradskogo Otdelenija Matematicheskogo Instituta, in V.A. Steklova Akademii Nauk SSSR p. 138-149

LI, L.L.(1991). Some sufficient conditions for the solvability of inverse eigenvalue problem. **Linear Algebra and Applications**. n 148. pp 225-236.

LI, L.L.(1995). Sufficient conditions for the solvability of inverse eigenvalue problem. **Linear Algebra and Applications**. n 221. pp 117-129.

MARTINEZ, J. M.; SANTOS, S. A. (1995). **Métodos computacionais de otimização**. In: Colóquio Brasileiro de Matemática, 20, Rio de Janeiro: Instituto de Matemática Pura e Aplicada . 24-28 julho, 256 pp.

MIERLO, J. van. (1981). Second order design: precision and reliability aspects. **Allgemeine Vermessung-Nachrichten**, Karlsruhe, v. 88, n. 3, p. 95-105.

NINKOV, T. ; SCHMITT, G. Eine Methode Gewichtsoptimierung in geodätischen Netzen. **Allgemeine Vermessungs-Nachrichten**, Karlsruhe, v. 90, n. 6 , 1983, p. 216-222.

NOCEDAL, J.; OVERTON, M. L. (1983). **Numerical methods for solving inverse eigenvalue problems**. Lectures Notes in Mathematic. n. 1005, ed. Springer-Verlag, New York-Berlin, p.212-226.

NOCEDAL, J.; WRIGHT, S.T. (1999). **Numerical optimization**. Springer series in operations recherche. Springer-Verlag New York. 635 pp

PELZER, H. (1980). Some criteria for the accuracy and the reability of networks. **Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften**, München, série B, p.55-67.

OLIVEIRA, R. (2003). **Otimização dos pesos das observações geodésicas pelo problema de valor próprio inverso**. Curitiba, 2003. Dissertação de Mestrado. Curso de Pós-Graduação em Ciências Geodésicas da Universidade Federal do Paraná, 95 pp.

RAO, C. R. ; MITRA, S. K. (1971) **Generalized inverse of matrices and its applications**. New York. J. Wiley.

SÁ, C. C. P. (1985). **Otimização de observações em redes geodésicas horizontais**. Rio de Janeiro, Dissertação (Mestrado em Ciências em Engenharia de Sistemas) – Instituto Militar de Engenharia.

SCHMITT, G. (1997). Spectral analysis and optimization of two dimensional networks. **Geomatics Research Australasia**. n. 67, p. 47-64.

SCHREIBER, O. (1882). Anordnung der winkelbeobachtung im Gottinger Basisnetz. **Zeitschrift Für Vermessungswesen**. v. 11. pp. 129-161.

SCHWIEGER, V. (2001). Zur Konstruktion synthetischer Kovarianzmatrizen. **Zeitschrift für Vermessungswesen**, Stuttgart. n. 3, p. 143-150.

STOPAR, B. (2001). Second order design of horizontal GPS net. **Survey Review**, v. 36, n. 279, p. 44-53.

Sun,J.G.; YE, Q. (1986). The unsolvability of inverse algebraic eigenvalue problems almost everywhere. **Journal Computation Mathematic**. n 4 pp 212-226.

Sun,J.G.; (1986). The unsolvability of multiplicative inverse eigenvalue problems almost everywhere. **Journal Computation Mathematic**. n 4 pp 227-244.

Xu, S.F. (1992). On the necessary conditions for the solvability of algebraic inverse eigenvalue problems. **Journal Computations Mathematic**. n 10 pp 93-97.

Xu, S.F. (1992). On the sufficient conditions for the solvability of algebraic inverse eigenvalue problems. **Journal Computations Mathematic**. n 10 pp 93-97.

VAN MIERLO, J. (1981). Second order design: precision an reliability aspects. **Allgemeine Vermessung-Nachrichten**, Karlsruhe, v. 91, n. 8, p. 27-33.

WELSCH, W. ;HEUNECKE, O. ;KUHLMANN, H. (2000). **Auswertung geodätischer Überwachungsmessungen**. Heidelberg: Wichmann. (Handbuch Ingenieurgeodäsie).

WOLF, H. (1961). Zur Kritik von Schereibers Satz uber die Gewichtsverteilung in Basisnetzen. **Zeitschrift Für Vermessungsweswn**. v. 86. pp. 177-179

WANG, S.-G. & CHOW, S.-C. (1994). **Advanced Linear Models: Theory and Applications**. Dekker, New York.

**APÊNDICE A – MATRIZ DEFINIDA E SEMI DEFINIDA POSITIVA**

### A.1-MATRIZ DEFINIDA POSITIVA

Uma matriz  $\mathbf{B}$   $n \times n$  é definida positiva se o produto interno entre vetores  $\langle \mathbf{B}\mathbf{X}, \mathbf{X} \rangle$  for maior que zero para todo vetor  $\mathbf{X}$  de dimensão  $n \times 1$  e diferente de zero, ou seja,

$$\mathbf{X}^t \mathbf{B} \mathbf{X} > 0, \quad \mathbf{X} \neq \mathbf{0}. \quad (\text{A1})$$

Uma matriz  $\mathbf{B}$  definida positiva apresenta as seguintes propriedades:

- a)  $\det(\mathbf{B}) > 0$  o que implica que  $\mathbf{B}$  é não singular, ou seja, é inversível;
- b) os valores próprios de  $\mathbf{B}$  são reais e positivos;
- c) se  $\mathbf{B}$  é definida positiva então  $\mathbf{X}^t \mathbf{B} \mathbf{X} = \text{tr}(\mathbf{X} \mathbf{X}^t \mathbf{B})$ .

### A.2-CONDIÇÕES PARA UMA MATRIZ SER DEFINIDA POSITIVA

Cada um dos testes abaixo estipula condições necessárias e suficientes para que uma matriz  $\mathbf{B}$  simétrica seja definida positiva. Uma matriz  $\mathbf{B}$  simétrica será positiva definida se passar por qualquer um dos testes abaixo:

(C1) A matriz  $\mathbf{B}$  é definida positiva se e só se puder ser reduzida a uma matriz triangular superior utilizando as operações elementares sobre linhas (escalonamento) e se os elementos da diagonal da matriz resultante (pivôs) forem todos positivos.

(C2) Um menor principal de  $\mathbf{B}$  é o determinante de qualquer submatriz de  $\mathbf{B}$  obtida da eliminação das últimas  $k$  linhas e  $k$  colunas. A matriz  $\mathbf{B}$  é definida positiva se e só se todos os seus menores principais forem positivos.

(C3) A matriz  $\mathbf{B}$  é definida positiva se e só se todos os seus valores próprios forem positivos.

Os seguintes testes definem condições necessárias para que a matriz  $\mathbf{B} = [b_{ij}]_{n \times n}$  seja definida positiva.

(T1) Os elementos da diagonal da matriz  $\mathbf{B}$  são positivos.

(T2) O elemento da matriz  $\mathbf{B}$  com maior valor absoluto está na diagonal principal da matriz  $\mathbf{B}$ .

(T3)  $b_{ii} b_{jj} > |b_{ij}|^2$  ( $i \neq j$ ).

### A.3-MATRIZ SEMI-DEFINIDA POSITIVA

Uma matriz é semi-definida positiva se

$$\mathbf{X}^t \mathbf{B} \mathbf{X} \geq 0, \quad \mathbf{X} \neq \mathbf{0} \quad (\text{A2})$$

Se as desigualdades em (A1) e (A2) forem invertidas, então  $\mathbf{B}$  é definida negativa e semidefinida negativa, respectivamente. As matrizes definidas positivas (ou negativas) são inversíveis e suas inversas também são definidas positivas (ou negativas).

**APÊNDICE B-NORMAS**

## NORMAS DE VETORES

Uma norma de um vetor qualquer  $\mathbf{X}$  de dimensão finita, que é representada por  $\|\mathbf{X}\|$ , é uma função real que satisfaz as quatro condições seguintes, considerando  $\mathbf{X}$  e  $\mathbf{Y}$  com a mesma dimensão:

1.  $\|\mathbf{X}\| \geq 0$ ;
2.  $\|\mathbf{X}\| = 0$  se e só se  $\mathbf{X} = 0$ ;
3.  $\|r\mathbf{X}\| = r\|\mathbf{X}\|$  para qualquer escalar  $r$ .
4. vale a desigualdade triangular  $\|\mathbf{X} + \mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$

O comprimento de um vetor é fornecido pela sua norma. Considerando um vetor arbitrário  $\mathbf{X} = [x_1 \ x_2 \ \dots \ x_n]^t$ , algumas das normas mais comuns para este vetor são:

- Norma euclidiana ou norma  $l_2$ :  $\|\mathbf{X}\|_2 = \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle} = \sqrt{\mathbf{X}^t \mathbf{X}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ ;
- Norma  $l_1$ :  $\|\mathbf{X}\|_1 = |x_1| + |x_2| + \dots + |x_n|$ ;
- Norma  $l_\infty$ :  $\|\mathbf{X}\|_\infty = \max(|x_1|, |x_2|, \dots, |x_n|)$ ;
- Norma  $l_p$  ( $p \geq 1$ ):  $\|\mathbf{X}\|_p = \left( |x_1|^p + |x_2|^p + \dots + |x_n|^p \right)^{\frac{1}{p}}$

## VETORES NORMALIZADOS E DISTÂNCIA

Um vetor unitário é um vetor cuja norma, ou comprimento, é igual à unidade. Normaliza-se um vetor multiplicando-o pelo inverso de sua norma, em consequência disto vetores normalizados são vetores unitários. Um conjunto de vetores se diz ortonormal se os vetores deste conjunto forem ortogonais entre si e ainda cada um desses vetores for unitário.

A distância entre dois vetores  $\mathbf{X}$  e  $\mathbf{Y}$  é dada por  $\|\mathbf{X} - \mathbf{Y}\|$ . O valor desta distância e também a designação de um vetor como unitário estão relacionados à norma escolhida.

## NORMA DE MATRIZES

A norma de uma matriz quadrada  $\mathbf{B}$ , representada por  $\|\mathbf{B}\|$ , é a função real que satisfaz as seguintes condições para todas as matrizes  $\mathbf{B}$  e  $\mathbf{H}$ , de mesma dimensão

1.  $\|\mathbf{B}\| \geq 0$ ;
2.  $\|\mathbf{B}\| = 0$  se e só se  $\mathbf{B} = \mathbf{0}$ ;
3.  $\|r\mathbf{B}\| = r\|\mathbf{B}\|$  para qualquer escalar  $r$ ;
4. Desigualdade Triangular:  $\|\mathbf{B} + \mathbf{H}\| \leq \|\mathbf{B}\| + \|\mathbf{H}\|$
5. Condição de Compatibilidade:  $\|\mathbf{B}\mathbf{H}\| \leq \|\mathbf{B}\|\|\mathbf{H}\|$

Devido à condição de compatibilidade, nem todas as normas de vetores podem ser generalizadas à norma de matrizes. Duas normas de vetores que são extensíveis à norma de matrizes são as normas  $l_1$ , representada no caso de norma de matriz por  $L_1$  e a norma euclidiana ou de Frobenius.

A norma  $L_1$  é dada por,

$$\|\mathbf{B}\|_1 = \max_{j=1,2,\dots,n} \left( \sum_{i=1}^n |b_{ij}| \right) \quad (\text{B1})$$

que é a maior das somas dos valores absolutos dos elementos de cada coluna.

A norma euclidiana ou de Frobenius, para uma matriz  $\mathbf{B} = [b_{ij}]_{n \times n}$ , é dada por,

$$\|\mathbf{B}\|_F = \left( \sum_{i=1}^n \sum_{j=1}^n |b_{ij}|^2 \right)^{\frac{1}{2}}. \quad (\text{B2})$$