

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

Vinícius Sousa Fazio

**INTERPOLAÇÃO ESPACIAL:
UMA COMPARAÇÃO ANALÍTICA ENTRE REDES RBF E
KRIGAGEM**

Florianópolis(SC)

2013

Vinícius Sousa Fazio

**INTERPOLAÇÃO ESPACIAL: UMA COMPARAÇÃO ANALÍTICA
ENTRE REDES RBF E KRIGAGEM**

Dissertação submetida ao Programa de Pós-
Graduação em Ciências da Computação para
a obtenção do Grau de Mestrado.
Orientador: Mauro Roisenberg, Dr.

Florianópolis(SC)

2013

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Fazio, Vinícius Sousa

Interpolação Espacial : Uma Comparação Analítica entre
Redes RBF e Krigagem / Vinícius Sousa Fazio ; orientador,
Mauro Roisenberg - Florianópolis, SC, 2013.

109 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Ciência da Computação.

Inclui referências

1. Ciência da Computação. 2. Interpolação Espacial. 3.
Redes RBF. 4. Krigagem. I. Roisenberg, Mauro. II.
Universidade Federal de Santa Catarina. Programa de Pós-
Graduação em Ciência da Computação. III. Título.

Vinícius Sousa Fazio

**INTERPOLAÇÃO ESPACIAL: UMA COMPARAÇÃO ANALÍTICA
ENTRE REDES RBF E KRIGAGEM**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestrado”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciências da Computação.

Florianópolis(SC), 06 de maio 2013.

Ronaldo dos Santos Mello, Dr.
Coordenador

Mauro Roisenberg, Dr.
Orientador

Banca Examinadora:

Mauro Roisenberg, Dr.
Presidente

Paulo Martins Engel, Dr.

Carlos Antonio Oliveira Vieira, Dr.

Pedro Alberto Barbeta, Dr.

Sílvia Modesto Nassar, Dr.

AGRADECIMENTOS

Agradeço ao orientador Mauro Roisenberg por sempre manter um olhar crítico sobre o trabalho, além de toda ajuda. Agradeço ao pessoal do laboratório L3C, especialmente aos colegas Márcio Weck, Leandro Figueiredo e Cássio Conti, pelas muitas e valiosas discussões. Agradeço a todos conhecidos, amigos e familiares que me apoiaram.

*Reality is that which when you stop believing
in it, it doesn't go away.*

Philip K. Dick

RESUMO

Entre diversos métodos de interpolação espacial, dois dos mais populares são a Krigagem e as Redes de Função de Base Radial (Redes RBF). Este trabalho desenvolve uma comparação analítica entre os métodos e demonstra matematicamente que ambos são equivalentes e produzem resultados idênticos quando configurados apropriadamente. É demonstrado que os dois métodos possuem uma estrutura semelhante e, como consequência, características da Krigagem, tais como variância da interpolação, também podem ser implementadas em Redes RBF, desde que sejam respeitadas as restrições exigidas por cada método. Apesar da equivalência matemática, é concluído pela complexidade computacional dos dois algoritmos que as Redes RBF são bem mais rápidas que a Krigagem porque as Redes RBF interpolam de uma forma em que os pesos não precisam ser recalculados para cada ponto interpolado, necessitando de menos processamento. Foram feitos experimentos para mostrar na prática os resultados teóricos obtidos. Alguns softwares comerciais e gratuitos que implementam a Krigagem foram testados a fim de verificar se suas implementações possuem a complexidade computacional do algoritmo original ou se otimizações foram implementadas.

Palavras-chave: Krigagem. Redes RBF. Redes Neurais. Comparação. Interpolação Espacial. Redes de Função de Base Radial. RBF Networks. Kriging

ABSTRACT

Between several spatial interpolation methods, two of the most popular are Kriging and Radial Basis Function Networks (RBF Network). This study develops an analytical comparison between both methods and shows mathematically that they are equivalent if properly configured. It is shown that both methods have a similar structure and, as consequence, features of Kriging, such as interpolation variance, can be implemented in RBF Networks, considering the constraints demanded by each method. Complexity was calculated for both methods to show the relative speed of RBF Networks. It is shown that both methods share a common structure and, as a consequence, all improvements in one method can be implemented in RBF Networks. Besides the mathematical equivalence, it is concluded by the computational complexity of both algorithms that RBF Networks are faster than Kriging because RBF Networks interpolates in a way that the weights do not need to be recalculated for each interpolated point, saving processing time. Experiments were made to show in practice the theoretical results obtained. Some free and commercial softwares that implement Kriging were tested to verify if their implementations have the same complexity of the original algorithm or if they implemented optimizations.

Keywords: Kriging. RBF Networks. Neural Networks. Comparison. Spatial Interpolation. Radial Basis Function Networks.

LISTA DE FIGURAS

Figura 1	Exemplo de interpolação espacial	25
Figura 2	Exemplo de aplicação em meteorologia	26
Figura 3	Exemplo de aplicação em oceanografia	27
Figura 4	Efeito não desejado da Krigagem Simples quando a tendência não é zero	41
Figura 5	Krigagem Ordinária vs Krigagem Simples	45
Figura 6	Exemplo de semivariograma experimental	58
Figura 7	Parâmetros do modelo de semivariograma.	59
Figura 8	Alguns exemplos gráficos de modelos de semivariogramas. . .	60
Figura 9	Comportamento da Krigagem variando seus parâmetros. . . .	61
Figura 10	Exemplo de dados anisotrópicos.	65
Figura 11	Exemplo de um modelo de semivariograma isotrópico e anisotrópico.	67
Figura 12	Exemplo de um modelo de semivariograma anisotrópico e seus parâmetros.	68
Figura 13	Exemplo de um modelo de covariograma comparado ao seu semivariograma equivalente.	69
Figura 14	Alguns exemplos gráficos de modelos de semivariogramas. . .	70
Figura 15	Exemplo de diagrama de uma Rede RBF	71
Figura 16	Exemplo da interpolação RBF com somente uma entrada. . . .	71
Figura 17	Exemplo de diagrama de uma Rede RBF com <i>bias</i>	74
Figura 18	Exemplo de funções de ativação com duas dimensões.	76
Figura 19	Exemplo de interpolação RBF com dados anisotrópicos. . . .	77
Figura 20	Parâmetros da função de ativação anisotrópica.	78
Figura 21	Experimento usando 50 dados amostrais de volume de minério de ferro.	96
Figura 22	Experimento usando 467 dados amostrais de precipitação. . .	96

LISTA DE TABELAS

Tabela 1	Experimento de aumento de dados amostrais	97
Tabela 2	Experimento de aumento de pontos interpolados	97
Tabela 3	Experimento com software comercial	98
Tabela 4	Experimento de aumento de dados amostrais no mGstat	99

LISTA DE ABREVIATURAS E SIGLAS

IDW	Inverse Distance Weighting
MLP	Multi-Layer Perceptron
RBF	Radial Basis Functions
SVM	Support Vector Machine

LISTA DE ALGORITMOS

1	Krigagem Simples	49
2	Krigagem Ordinária	50
3	Krigagem Universal	51
4	Redes RBF Exatas sem <i>bias</i>	79
5	Redes RBF Exatas com <i>bias</i> constante	80
6	Redes RBF Exatas com <i>bias</i> variável	81

SUMÁRIO

1 INTRODUÇÃO	25
1.1 JUSTIFICATIVA	28
1.2 OBJETIVO PRINCIPAL	28
1.3 OBJETIVOS ESPECÍFICOS	28
1.4 LIMITAÇÕES	29
1.5 ORGANIZAÇÃO	29
2 MÉTODOS DE INTERPOLAÇÃO ESPACIAL	31
2.1 KRIGAGEM	32
2.1.1 Krigagem Simples	38
2.1.2 Krigagem Ordinária	40
2.1.3 Krigagem Universal	44
2.1.4 Algoritmo da Krigagem	49
2.1.5 Variância da Krigagem	52
2.1.5.1 Variância na Krigagem Simples	52
2.1.5.2 Variância na Krigagem Ordinária	53
2.1.5.3 Variância na Krigagem Universal	55
2.1.6 Semivariograma	57
2.1.6.1 Semivariograma experimental	57
2.1.6.2 Modelo de semivariograma	58
2.1.6.3 Validade do semivariograma	60
2.1.6.4 Ajuste do semivariograma	63
2.1.6.5 Anisotropia em semivariograma	64
2.1.7 Covariograma	66
2.2 REDES RBF	69
2.2.1 Treinamento	72
2.2.2 Variações	73
2.2.3 Anisotropia em Redes RBF	76
2.2.4 Algoritmo das Redes RBF Exatas	79
2.3 KRIGAGEM NEURAL E COMPARAÇÕES	82
3 COMPARAÇÃO ANALÍTICA	85
3.1 COMPARAÇÃO DOS MÉTODOS DE INTERPOLAÇÃO	85
3.2 COMPARAÇÃO ENTRE OS PARÂMETROS	89
3.3 COMPARAÇÃO DE COMPLEXIDADE	93
4 EXPERIMENTOS	95
4.1 EXPERIMENTO DE AUMENTO DE DADOS AMOSTRAIS ...	95
4.2 EXPERIMENTO COM AUMENTO DE PONTOS INTERPOLADOS	97

4.3 EXPERIMENTOS COM SOFTWARE COMERCIAL	98
4.4 EXPERIMENTOS COM BIBLIOTECA DE KRIGAGEM.....	98
5 CONCLUSÃO.....	101
5.1 TRABALHOS FUTUROS.....	102
Referências Bibliográficas	105

1 INTRODUÇÃO

Muitas áreas precisam utilizar dados contínuos distribuídos em um mapa ou superfície para tomar decisões e realizar planejamentos. Alguns exemplos destes mapas são: mapas de precipitação de uma região, mapas de volume de minérios em uma jazida, mapas de poluição de uma cidade, entre diversos outros. Porém, muitas vezes é muito difícil obter tantos dados para gerar um mapa por diversas razões, entre elas, o elevado custo ou a dificuldade de realização das medições (LI; HEAP, 2008).

A interpolação espacial é o processo de gerar mapas de dados a partir de uma pequena amostra de medições, de forma que este mapa seja uma representação relativamente confiável da realidade, como mostrado na Figura 1. Uma medição é um valor que representa uma variável, como quantidade de precipitação, e uma posição, como as coordenadas horizontais e verticais de um mapa de duas dimensões. O processo de interpolação espacial encontra uma série de aplicações em áreas como geologia, mineração, hidrologia e meteorologia (MATÍAS et al., 2004).

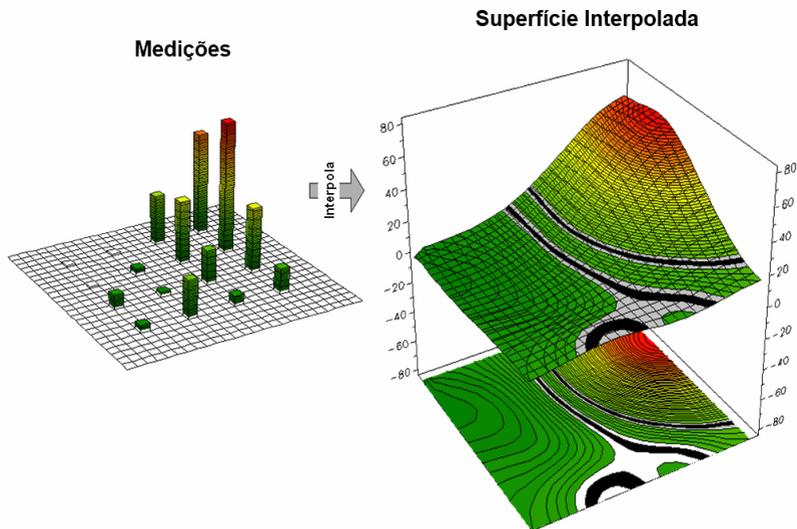


Figura 1: Exemplo de interpolação espacial (BERRY, 2007)

A Figura 2 mostra um exemplo de interpolação espacial na área de

meteorologia. Os modelos de previsão do tempo precisam de dados amostrais regularmente espaçados em forma de grade. Porém, as estações que medem informações meteorológicas são distribuídas não uniformemente, e com muito menos pontos que o exigido pelos modelos de previsão, por causa de dificuldades geográficas e financeiras de se ter medições em cada ponto da grade. Por isso, a interpolação espacial interpola as medições em toda a grade de previsão do tempo.

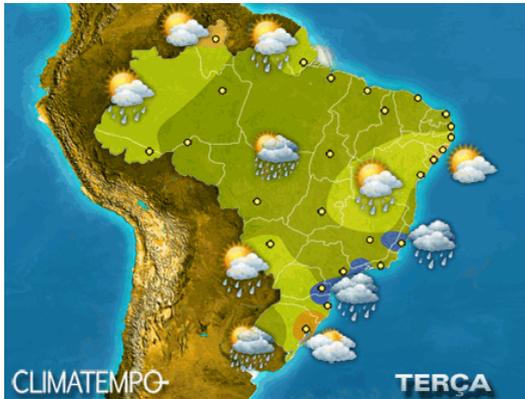


Figura 2: Exemplo de aplicação em meteorologia (CLIMATEMPO, 2013)

A Figura 3 mostra um exemplo de interpolação espacial na área de oceanografia. É possível modelar dados de altura da maré e direção a partir de alguns dados amostrais.

Existem diversos métodos para realizar interpolação espacial, cada um com suas características, vantagens e desvantagens. Este trabalho foca na comparação de dois métodos: a Krigagem e as Redes Neurais Artificiais.

A Krigagem realiza a interpolação através da análise da variação dos dados amostrais associado à distância entre eles. É um método com forte base estatística, da área de Geoestatística, e que utiliza o *Best Linear Unbiased Estimator* (melhor estimador linear não viesado). Diversos estudos foram feitos com o objetivo de comparar o erro da interpolação espacial obtida através da Krigagem em relação a outros métodos (SAKATA; ASHIDA; ZAKO, 2010; MATÍAS et al., 2004; BADEL; ANGORANI; PANAHI, 2011; COSTA; PRONZATO; THIERRY, 1999). Entretanto, a Krigagem possui a desvantagem de ser computacionalmente custosa para conjuntos grandes de dados amostrais.

Entre as técnicas utilizadas para interpolação, uma das mais conhecidas são as Redes Neurais Artificiais, que tentam imitar a forma que o cérebro

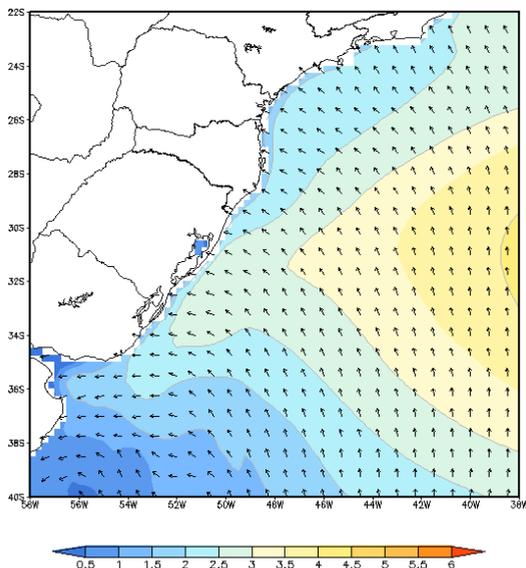


Figura 3: Exemplo de aplicação em oceanografia

bro humano realiza determinadas tarefas para resolver problemas complexos, como reconhecimento de padrões, interpolação e classificação (HAYKIN, 2008; BISHOP, 1996). As Redes Neurais Artificiais possuem a vantagem de serem facilmente utilizáveis, sem necessitar conhecimento prévio do domínio do problema. As Redes de Função de Base Radial, conhecidas como Redes RBF, são Redes Neurais Artificiais que utilizam a distância entre os pontos para a interpolação, assim como a Krigagem.

Existem diversos trabalhos que descrevem comparações entre a Krigagem e as Redes Neurais Artificiais em aplicações de Interpolação Espacial. Alguns trabalhos (SATHYANARAYANAMURTHY; CHINNAM, 2009; ZHANG; GOVINDARAJU, 2000; COSTA; PRONZATO; THIERRY, 1999) apenas comparam qual método tem o melhor desempenho, através de experimentos com conjuntos de dados de teste, sem analisar as diferenças e similaridades dos algoritmos e parâmetros. Isso leva a resultados que classificam métodos em melhores ou piores somente para o conjunto de dados utilizados, que é limitante. Analisar as características que um método tem de diferente do outro para realizar a interpolação pode levar a uma melhor comparação para verificar se um método é superior ao outro de forma mais geral.

Outros trabalhos tentam mesclar os métodos de Redes Neurais Artifi-

ciais e Krigagem. Essa técnica normalmente é chamada de Krigagem Neural (LIN; CHEN, 2004; KOIKE; MATSUDA; GU, 2001; DEMYANOV et al., 1998; ZHENG et al., 2009; DEMYANOV et al., 2001). Dessa forma, um novo método de interpolação é criado utilizando características dos métodos originais. Nestes trabalhos, a conclusão normalmente obtida é que o novo método é melhor que os métodos originais, entretanto, por se limitar a um conjunto de dados específicos, estas propostas não podem generalizar suas conclusões para qualquer conjunto de dados.

1.1 JUSTIFICATIVA

Em função do que foi discutido anteriormente, percebe-se a quase inexistência de trabalhos que realizem uma comparação analítica, que identifique nos algoritmos o que há em comum e de diferente entre os dois métodos. Na Krigagem são utilizadas as propriedades estocásticas dos dados amostrais, enquanto as Redes RBF são mais eficientes computacionalmente. Dessa forma, considerando as restrições de cada método, pode ser possível que as melhorias feitas em um dos métodos possam ser utilizadas no outro método, fazendo com que áreas distintas como Geoestatística e Inteligência Artificial possam colaborar entre si.

1.2 OBJETIVO PRINCIPAL

O objetivo principal deste trabalho é fazer uma comparação analítica entre os métodos de Krigagem e de Redes RBF para realização da interpolação espacial.

1.3 OBJETIVOS ESPECÍFICOS

Para alcançar o objetivo principal, os seguintes objetivos específicos deverão ser atingidos:

- Fazer um levantamento bibliográfico dos métodos analisados, verificando as comparações já realizadas;
- Comparar os métodos procurando as semelhanças e diferenças nos seus algoritmos;
- Realizar uma comparação de complexidade dos métodos, através da notação “Big O”;

- Realizar experimentos para verificar o desempenho calculado pela complexidade;
- Realizar experimentos em softwares comerciais e bibliotecas gratuitas, e compará-los com os resultados obtidos;
- Propor uma melhoria visando mesclar as qualidades de ambos os métodos.

1.4 LIMITAÇÕES

O escopo do trabalho foi comparar as seguintes variações de Krigagem e Redes RBF:

- Krigagem Simples, Krigagem Ordinária e Krigagem Universal;
- Redes RBF exatas, que utilizam uma função de base radial do neurônio para cada ponto disponível do conjunto de treinamento, com e sem *bias*.

É possível ampliar a comparação para outros tipos de Krigagem e Redes RBF, mas não foram analisadas nesta dissertação por haver inúmeras variações.

1.5 ORGANIZAÇÃO

Esta dissertação está organizada nos seguintes capítulos:

- **Capítulo 2 - Métodos de Interpolação Espacial:** É feito um levantamento bibliográfico dos métodos analisados, buscando comparações já realizadas. Além disso, os métodos são descritos detalhadamente;
- **Capítulo 3 - Comparação Analítica:** É feita a comparação dos algoritmos de cada método e calculadas as suas complexidades;
- **Capítulo 4 - Experimentos:** Contém experimentos que usam os métodos analisados para verificar os resultados da comparação analítica;
- **Capítulo 5 - Conclusão:** É mostrada a conclusão obtida pelo trabalho.

2 MÉTODOS DE INTERPOLAÇÃO ESPACIAL

O objetivo da interpolação espacial é estimar valores para regiões do espaço utilizando medições feitas em locais próximos. Existem muitos métodos diferentes para realizar a interpolação espacial. Alguns exemplos:

- **Vizinho mais Próximo:** Também conhecido como *Nearest Neighbour*. Estima o valor desconhecido como igual à medição mais próxima conhecida (LI; HEAP, 2008);
- **Ponderação pelo Inverso da Distância (IDW):** Mais conhecido como *Inverse Distance Weighting*. Esse método estima o valor desconhecido como igual à média das medições ponderadas pelo inverso da distância do ponto a ser estimado. Ela se diferencia da Krigagem por ter pesos que variam proporcionais à distância sem levar em conta o quanto as variações das distâncias influenciam no comportamento dos dados (LI; HEAP, 2008);
- **Splines:** Estima o valor desconhecido como igual a um polinômio que utiliza somente as medições mais próximas (LI; HEAP, 2008);
- **Séries de Fourier:** Estima o valor desconhecido utilizando funções de seno e cosseno. Exige que os dados sejam periódicos (LI; HEAP, 2008);
- **Krigagem:** Estima o valor desconhecido utilizando um modelo da variação dos dados em função da distância, conhecido como semivariograma (CRESSIE, 1993; STEIN; MEER; GORTE, 2002; LI; HEAP, 2008). Este método é descrito em detalhes na Seção 2.1;
- **Redes de Funções de Base Radial (RBF):** Estima o valor desconhecido utilizando neurônios que usam funções de ativação de base radial, como a função Gaussiana (BISHOP, 1996; HAYKIN, 2008). Este método é descrito em detalhes na Seção 2.2;
- **Redes Perceptron Multi-Camadas (MLP):** Outra rede neural, mais conhecida por *Multi-Layer Perceptron*, que utiliza conceitos diferentes das Redes RBF, mas também inspirada no cérebro humano. A grande diferença com as Redes RBF é a ausência de funções radiais com centros que são ativados de acordo com a proximidade dos dados de entrada. Ao invés disso, as entradas são multiplicadas por pesos associados a cada neurônio e depois passam por uma função de transferência

que determina a saída do neurônio. Esse procedimento pode ser encaixado em várias camadas de neurônios até serem combinados em uma saída (BISHOP, 1996);

- **Markov Bayes:** Método que utiliza probabilidades condicionais para interpolar (SIL; SRINIVASAN, 2009).

Todos os métodos descritos neste trabalho utilizam a mesma representação dos dados. Cada dado possui um ponto associado. O ponto, por sua vez, possui uma coordenada horizontal e vertical, que representam a sua posição em um mapa de duas dimensões. Para representá-las, utiliza-se a notação x_a , para a coordenada horizontal do ponto a , e y_a , para a coordenada vertical do ponto a . Além disso, cada dado possui o valor da variável que o dado está representando, como quantidade de precipitação, volume de minério, temperatura, entre outros, que varia conforme a aplicação. O valor da variável é representado pela letra z , enquanto a distinção do ponto fica subscrito, como z_a , que representa o valor do ponto a .

2.1 KRIGAGEM

A Krigagem é um método fortemente baseado em estatística (CRESSIE, 1993; CHAUVET; GALLI, 1982; VENKATRAM, 1988). Ela foi criada por Danie G. Krige e desenvolvida mais tarde por Georges Matheron. O objetivo do método era estimar a distribuição de minérios em uma região da África do Sul.

Aqui será descrito detalhadamente o método da Krigagem, desde o que é assumido até a dedução de todas as fórmulas utilizadas pelo método. Isso se deve a dificuldade de se encontrar uma referência que faça essa descrição de forma didática e completa.

A Krigagem utiliza os dados amostrais e a relação espacial que os dados têm entre si, que é a variação quadrática esperada dada uma distância. Essa relação é chamada de semivariograma. Além disso, a Krigagem leva em consideração o valor médio dos dados, chamado de tendência. Utilizando essas duas informações, tendência e semivariograma, a Krigagem realiza a interpolação através de uma média ponderada dos dados amostrais de forma que o erro esperado seja minimizado.

A Krigagem divide o dado em duas partes: uma parte chamada de tendência, que é o valor esperado ou valor médio da interpolação, e outra chamada de ruído, que é uma parte imprevisível, cuja média é igual a zero. Na equação a seguir, a tendência é representada por t , o ruído por s e o dado

por z (CRESSIE, 1993):

$$z = s + t, \quad (2.1.1)$$

onde o valor esperado do ruído é zero:

$$E(s) = 0, \quad (2.1.2)$$

e que é possível concluir que o valor esperado dos dados é igual à tendência:

$$\begin{aligned} E(z) &= E(s) + E(t) \\ &= 0 + E(t) \\ &= E(t) \end{aligned} \quad (2.1.3)$$

Para interpolar, a Krigagem realiza uma soma ponderada dos dados amostrais:

$$z_j = \sum_i \omega_i z_i, \quad (2.1.4)$$

sendo que i indica os pontos amostrais e j indica os pontos com valores desconhecidos que precisam ser interpolados. O peso de cada dado amostral usado é representado pelo ω .

A Krigagem é um método não viesado, o que significa que o valor interpolado esperado é igual ao valor real esperado. Portanto:

$$\begin{aligned} E(z_j) &= E\left(\sum_i \omega_i z_i\right) \\ &= \sum_i \omega_i E(z_i) \end{aligned} \quad (2.1.5)$$

Os pesos são calculados de forma que o valor esperado do erro quadrático da interpolação seja o menor possível. A Krigagem utiliza o *Best Linear Unbiased Estimator* (melhor estimador linear não viesado) para calcular os pesos de forma determinística. A equação do erro médio quadrático é:

$$\sigma^2 = E\left(\overbrace{\sum_i \omega_i z_i}^{\text{Krigagem}} - z_j\right)^2, \quad (2.1.6)$$

ou seja, o valor esperado do quadrado da diferença entre o valor interpolado e o seu valor real, que é representado por σ^2 .

Aqui é introduzido um novo conceito usado na Krigagem: o semivari-

ograma. Ele é definido por:

$$\gamma(h) = \frac{E(z_a - z_b)^2}{2}, \quad (2.1.7)$$

que é a metade da diferença quadrática esperada entre os valores de dois pontos a e b cuja distância entre si é h . Mais detalhes sobre o semivariograma será tratado na Seção 2.1.6.

A função do erro, Equação (2.1.6), será reescrita para utilizar o semivariograma. Para isso, abre-se a equação:

$$\sigma^2 = E \left(\overbrace{\sum_{i1} \omega_{i1} z_{i1}}^A \overbrace{\sum_{i2} \omega_{i2} z_{i2}}^B - 2z_j \sum_i \omega_i z_i + z_j^2} \right) \quad (2.1.8)$$

A parte A da Equação (2.1.8) pode ser reescrita da seguinte forma:

$$\begin{aligned} A &= \sum_{i1} \omega_{i1} z_{i1} \sum_{i2} \omega_{i2} z_{i2} \\ &= \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i2} z_{i1} \end{aligned} \quad (2.1.9)$$

A multiplicação $z_{i1} z_{i2}$ é substituída pela equação algebricamente equivalente:

$$\begin{aligned} A &= \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \frac{1}{2} (z_{i1}^2 + z_{i2}^2 - (z_{i1} - z_{i2})^2) \\ &= \frac{1}{2} \left(\sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i1}^2 + \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i2}^2 - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} (z_{i1} - z_{i2})^2 \right) \end{aligned} \quad (2.1.10)$$

Sabe-se que:

$$\sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i1}^2 = \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i2}^2, \quad (2.1.11)$$

que é usado para simplificar a expressão A em:

$$\begin{aligned} A &= \frac{1}{2} \left(2 \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} z_{i1}^2 - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} (z_{i1} - z_{i2})^2 \right) \\ &= \frac{1}{2} \left(2 \sum_{i1} \omega_{i1} z_{i1}^2 \sum_{i2} \omega_{i2} - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} (z_{i1} - z_{i2})^2 \right) \end{aligned} \quad (2.1.12)$$

O somatório $\sum_{i2} \omega_{i2}$ pode ser substituído, para simplificar, por uma constante desconhecida c , que representa a soma dos pesos:

$$\begin{aligned} A &= \frac{1}{2} \left(2 \sum_i \omega_i z_i^2 c - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} (z_{i1} - z_{i2})^2 \right) \\ &= c \sum_i \omega_i z_i^2 - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \frac{1}{2} (z_{i1} - z_{i2})^2 \end{aligned} \quad (2.1.13)$$

Agora é reescrita a parte B da Equação (2.1.8):

$$\begin{aligned} B &= -2z_j \sum_i \omega_i z_i + z_j^2 \\ &= -2 \sum_i \omega_i z_i z_j + z_j^2 \end{aligned} \quad (2.1.14)$$

A multiplicação $z_i z_j$ é substituída pela equação algebricamente equivalente:

$$\begin{aligned} B &= -2 \sum_i \omega_i \frac{1}{2} (z_i^2 + z_j^2 - (z_i - z_j)^2) + z_j^2 \\ &= - \sum_i \omega_i (z_i^2 + z_j^2 - (z_i - z_j)^2) + z_j^2 \\ &= - \sum_i \omega_i z_i^2 - \sum_i \omega_i z_j^2 + \sum_i \omega_i (z_i - z_j)^2 + z_j^2 \\ &= - \sum_i \omega_i z_i^2 - z_j^2 \sum_i \omega_i + \sum_i \omega_i (z_i - z_j)^2 + z_j^2 \end{aligned} \quad (2.1.15)$$

Como já citado anteriormente, o somatório $\sum_{i1} \omega_{i1}$ pode ser substituído, para simplificar, por uma constante desconhecida c , que representa a

soma dos pesos:

$$\begin{aligned} B &= -\sum_i \omega_i z_i^2 - z_j^2 c + \sum_i \omega_i (z_i - z_j)^2 + z_j^2 \\ &= -\sum_i \omega_i z_i^2 + \sum_i \omega_i (z_i - z_j)^2 + (1 - c)z_j^2 \end{aligned} \quad (2.1.16)$$

Somando a expressão A e a expressão B , obtém-se:

$$\begin{aligned} A + B &= (1 - c)z_j^2 + (c - 1) \sum_i \omega_i z_i^2 - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \frac{1}{2} (z_{i1} - z_{i2})^2 + \sum_i \omega_i (z_i - z_j)^2 \\ &= (1 - c)(z_j^2 - \sum_i \omega_i z_i^2) - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \frac{1}{2} (z_{i1} - z_{i2})^2 + \sum_i \omega_i (z_i - z_j)^2 \end{aligned} \quad (2.1.17)$$

Para simplificar a equação e remover o termo com a constante c , é assumido que $c = 1$:

$$\sum_i \omega_i = c = 1 \quad (2.1.18)$$

Como a expressão $A + B$ representa a função de erro σ^2 , o resultado é:

$$\sigma^2 = E \left(-\sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \frac{1}{2} (z_{i1} - z_{i2})^2 + \sum_i \omega_i (z_i - z_j)^2 \right) \quad (2.1.19)$$

Utiliza-se o semivariograma para simplificar a equação em:

$$\sigma^2 = E \left(-\sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2\gamma_{ij} \right), \quad (2.1.20)$$

onde $\gamma_{i1,i2} = \gamma(\text{Distância}(i1, i2))$ e $\gamma_{ij} = \gamma(\text{Distância}(i, j))$. Essa equação é uma forma diferente de escrever a função de erro, porém equivalente à Equação (2.1.6). Essa forma será usada adiante.

Na equação original do erro, Equação (2.1.6), é possível substituir os

dados z por suas partes de tendência e ruído, conforme a Equação (2.1.1):

$$\begin{aligned}
 \sigma^2 &= E \left(\sum_i \omega_i (t_i + s_i) - (t_j + s_j) \right)^2 \\
 &= E \left(\sum_i \omega_i t_i + \sum_i \omega_i s_i - t_j - s_j \right)^2 \\
 &= \left(E \left(\sum_i \omega_i t_i \right) + E \left(\sum_i \omega_i s_i \right) - E(t_j) - E(s_j) \right)^2 \\
 &= \left(\sum_i \omega_i E(t_i) + E \left(\sum_i \omega_i s_i \right) - E(t_j) - E(s_j) \right)^2 \quad (2.1.21)
 \end{aligned}$$

Usa-se a Equação (2.1.3) para constatar que:

$$\begin{aligned}
 E(z_j) &= E(t_j) \\
 E(z_i) &= E(t_i) \\
 E(t_j) &= \sum_i \omega_i E(t_i), \quad (2.1.22)
 \end{aligned}$$

ou seja, o valor esperado da tendência real é igual à soma ponderada do valor esperado da tendência da amostra. Utiliza-se a Equação (2.1.22) para simplificar a equação em:

$$\begin{aligned}
 \sigma^2 &= \left(E \left(\sum_i \omega_i s_i \right) - E(s_j) \right)^2, \\
 &= E \left(\sum_i \omega_i s_i - s_j \right)^2, \quad (2.1.23)
 \end{aligned}$$

onde se conclui que a tendência não influencia a função de erro.

A função de erro (2.1.6) é uma equação de segundo grau. Para minimizá-la basta encontrar os pesos onde a derivada é igual a zero. Entretanto é necessário fazer isso de forma que a restrição da tendência, mostrada em (2.1.22), seja respeitada, mesmo que a função de erro não utilize a tendência, conforme mostrado em (2.1.23). Para isso, pode-se utilizar o Multiplicador de Lagrange (KLEIN, ; BERTSEKAS, 1982; MORDECAI, 2003). A Função de Lagrange

é definida por:

$$\Lambda(x, \lambda) = f(x) + \lambda g(x) \quad (2.1.24)$$

$$g(x) = c, \quad (2.1.25)$$

onde c é uma constante, $f(x)$ é a função a ser minimizada, $g(x)$ é a restrição e λ é uma variável extra, chamada de Multiplicador de Lagrange. A minimização respeitando a restrição é feita igualando a sua derivada a zero:

$$\nabla (f(x) + \lambda g(x)) = 0 \quad (2.1.26)$$

$$g(x) = c \quad (2.1.27)$$

Utiliza-se então o Multiplicador de Lagrange na Krigagem para minimizar a função de erro dada a restrição da tendência (CHAUVET; GALLI, 1982):

$$\nabla \left(\sigma^2 + \lambda_j (E(t_j) - \sum_i \omega_i E(t_i)) \right) = 0 \quad (2.1.28)$$

$$E(t_j) - \sum_i \omega_i E(t_i) = 0, \quad (2.1.29)$$

onde o Multiplicador de Lagrange é representado por λ_j , a tendência por t , a função do erro por σ^2 , os pesos por ω , os pontos amostrais por i e o ponto a ser interpolado por j .

Para continuar desenvolvendo as equações (2.1.28) e (2.1.29) e encontrar os pesos é necessário saber qual a tendência utilizada. Neste trabalho são analisados três tipos de Krigagem: Krigagem Simples, Krigagem Ordinária e Krigagem Universal. O que diferencia uma Krigagem da outra é como ela assume que seja a tendência.

2.1.1 Krigagem Simples

Na Krigagem Simples, a tendência, relacionada em (2.1.1), é uma constante conhecida e igual a zero (CRESSIE, 1993):

$$t = 0 \quad (2.1.30)$$

As equações (2.1.28) e (2.1.29) podem ser totalmente desenvolvidas

com a tendência assumida:

$$\nabla \left(\sigma^2 + \lambda_l (0 - \sum_i \omega_i 0) \right) = 0 \quad (2.1.31)$$

$$0 - \sum_i \omega_i 0 = 0, \quad (2.1.32)$$

que é simplificado em:

$$\nabla \sigma^2 = 0 \quad (2.1.33)$$

Agora, substitui-se o σ^2 da Equação (2.1.20):

$$\nabla \left(- \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i^2 \gamma_{ij} \right) = 0 \quad (2.1.34)$$

Após aplicar a derivada conforme o Multiplicador de Lagrange, as seguintes equações são geradas:

$$\begin{aligned} \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i^2 \gamma_{1j}}{\partial \omega_1} &= 0 \\ \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i^2 \gamma_{2j}}{\partial \omega_2} &= 0 \\ &\vdots \\ \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i^2 \gamma_{nj}}{\partial \omega_n} &= 0, \end{aligned} \quad (2.1.35)$$

onde n é o número de dados amostrais.

O número 2 acaba cancelado após derivar, além de vários termos que acabam zerados, resultando nas equações:

$$\begin{aligned} \sum_i \omega_i \gamma_{1i} &= \gamma_{1j} \\ \sum_i \omega_i \gamma_{2i} &= \gamma_{2j} \\ &\vdots \\ \sum_i \omega_i \gamma_{ni} &= \gamma_{nj}, \end{aligned} \quad (2.1.36)$$

que é equivalente à forma matricial:

$$\overbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{bmatrix}}^{\mathbf{\Gamma}} \overbrace{\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix}}^{\mathbf{\Omega}_k} = \overbrace{\begin{bmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{nj} \end{bmatrix}}^{\mathbf{Y}} \quad (2.1.37)$$

Os símbolos $\mathbf{\Gamma}$, $\mathbf{\Omega}_k$ e \mathbf{Y} estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz $\mathbf{\Gamma}$ e multiplicá-la pela matriz \mathbf{Y} :

$$\begin{aligned} \mathbf{\Gamma} \mathbf{\Omega}_k &= \mathbf{Y} \\ \mathbf{\Omega}_k &= \mathbf{\Gamma}^{-1} \mathbf{Y} \end{aligned} \quad (2.1.38)$$

Esse é o tipo de Krigagem mais simples, porém possui desempenho ruim quando a tendência dos dados não for zero. A Figura 4 mostra um exemplo de Krigagem Simples cuja média dos valores amostrais é zero e outro exemplo onde a Krigagem Simples foi feita utilizando os mesmos parâmetros, com a única diferença sendo a soma do valor 380 em todos os dados amostrais. O resultado ideal seria apenas uma mudança na escala dos dados, mas a interpolação e o formato deveriam ficar iguais, mas o que acontece é uma distorção na interpolação, por causa da assunção de tendência zero da Krigagem Simples. As cores representam os valores interpolados espacialmente, distribuídos nos eixos horizontal e vertical.

2.1.2 Krigagem Ordinária

Na Krigagem Ordinária, a tendência, relacionada em (2.1.1), é uma constante de valor desconhecido (CRESSIE, 1993):

$$t = k_c, \quad (2.1.39)$$

onde k_c é uma constante de valor desconhecido.

O processo descrito nesta seção é similar ao da Seção 2.1.1, já que a única mudança é a tendência.

As equações (2.1.28) e (2.1.29) podem ser totalmente desenvolvidas

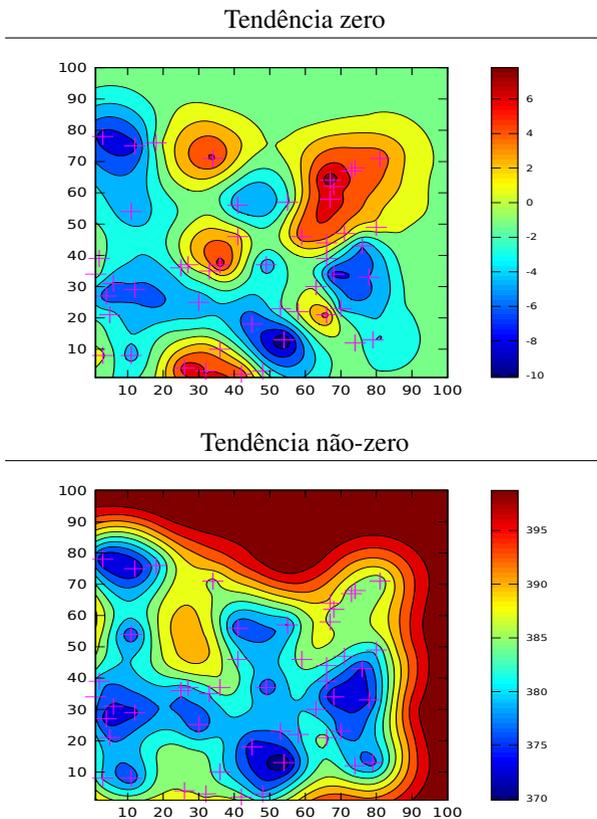


Figura 4: Efeito não desejado da Krigagem Simples quando a tendência não é zero. O resultado esperado era a mesma interpolação, mas com escala diferente.

com a tendência assumida. A Equação (2.1.28) pode ser escrita como:

$$\begin{aligned}\nabla \left(\sigma^2 + \lambda_l (k_c - \sum_i \omega_i k_c) \right) &= 0 \\ \nabla \left(\sigma^2 + \lambda_l k_c (1 - \sum_i \omega_i) \right) &= 0\end{aligned}\quad (2.1.40)$$

Como λ_l e k_c são constantes desconhecidas, a multiplicação $\lambda_l k_c$ será referenciada por λ_c para simplificar:

$$\nabla \left(\sigma^2 + \lambda_c (1 - \sum_i \omega_i) \right) = 0 \quad (2.1.41)$$

A Equação (2.1.29) pode ser escrita como:

$$\begin{aligned}k_c - \sum_i \omega_i k_c &= 0 \\ \sum_i \omega_i k_c &= k_c \\ \sum_i \omega_i &= 1\end{aligned}\quad (2.1.42)$$

A restrição $\sum_i \omega_i = 1$ já havia sido considerada nas equações para minimizar a variância, na Seção 2.1. Juntando as equações (2.1.41) e (2.1.42):

$$\nabla \left(\sigma^2 + \lambda_c (1 - \sum_i \omega_i) \right) = 0 \quad (2.1.43)$$

$$\sum_i \omega_i = 1 \quad (2.1.44)$$

Substituindo o σ^2 da Equação (2.1.20), as seguintes equações são geradas:

$$\nabla \left(-\sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i 2\gamma_{ij} + \lambda_c (1 - \sum_i \omega_i) \right) = 0 \quad (2.1.45)$$

$$\sum_i \omega_i = 1 \quad (2.1.46)$$

Após aplicar a derivada conforme o Multiplicador de Lagrange:

$$\begin{aligned}
 \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i 2\gamma_{1j} + \lambda_c - \lambda_c \sum_i \omega_i}{\partial \omega_1} &= 0 \\
 \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i 2\gamma_{2j} + \lambda_c - \lambda_c \sum_i \omega_i}{\partial \omega_2} &= 0 \\
 &\vdots \\
 \frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{1,i2} + \sum_i \omega_i 2\gamma_{nj} + \lambda_c - \lambda_c \sum_i \omega_i}{\partial \omega_n} &= 0 \\
 \sum_i \omega_i &= 1,
 \end{aligned} \tag{2.1.47}$$

onde n é o número de dados amostrais.

O número 2 acaba cancelado após derivar, além de vários termos que acabam zerados, resultando nas equações:

$$\begin{aligned}
 \sum_i \omega_i \gamma_{1i} + \lambda_c &= \gamma_{1j} \\
 \sum_i \omega_i \gamma_{2i} + \lambda_c &= \gamma_{2j} \\
 &\vdots \\
 \sum_i \omega_i \gamma_{ni} + \lambda_c &= \gamma_{nj} \\
 \sum_i \omega_i &= 1,
 \end{aligned} \tag{2.1.48}$$

que é equivalente à forma matricial:

$$\overbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}}^{\mathbf{\Gamma}} \overbrace{\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \\ \lambda_c \end{bmatrix}}^{\mathbf{\Omega}_k} = \overbrace{\begin{bmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{nj} \\ 1 \end{bmatrix}}^{\mathbf{Y}} \tag{2.1.49}$$

Os símbolos $\mathbf{\Gamma}$, $\mathbf{\Omega}_k$ e \mathbf{Y} estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz $\mathbf{\Gamma}$ e

multiplicá-la pela matriz Υ :

$$\begin{aligned}\Gamma \Omega_k &= \Upsilon \\ \Omega_k &= \Gamma^{-1} \Upsilon\end{aligned}\quad (2.1.50)$$

A Krigagem Ordinária é bastante utilizada por se aplicar na maioria dos problemas, já que a maioria dos problemas possuem tendência com valor desconhecido. Além disso, é utilizada apenas uma variável extra para isso, λ_c , o que torna possível sua utilização com poucos dados amostrais.

No exemplo mostrado na seção anterior, a Krigagem Simples obteve resultados diferentes quando somente a tendência dos dados variava. Na Krigagem Ordinária, foi feito o mesmo experimento, mas, como esperado, o resultado foi exatamente o mesmo, variando somente a escala, conforme mostrado na Figura 5.

2.1.3 Krigagem Universal

Na Krigagem Universal, a tendência, relacionada em (2.1.1), é um polinômio. Pode ser, por exemplo, um polinômio que utiliza as coordenadas do ponto (CRESSIE, 1993):

$$t = k_x x + k_y y + k_c, \quad (2.1.51)$$

onde k_x , k_y e k_c são os coeficientes do polinômio, e x e y são, respectivamente, as coordenadas horizontais e verticais do ponto onde a tendência é calculada.

Também aqui o processo é similar ao da Seção 2.1.2, já que a única mudança é a tendência.

As equações (2.1.28) e (2.1.29) podem ser totalmente desenvolvidas com a tendência assumida. A Equação (2.1.28) pode ser escrita como:

$$\begin{aligned}\nabla \left(\sigma^2 + \lambda_l \left(k_x x_j + k_y y_j + k_c - \sum_i \omega_i (k_x x_i + k_y y_i + k_c) \right) \right) &= 0 \\ \nabla \left(\sigma^2 + \left(\lambda_l k_x x_j + \lambda_l k_y y_j + \lambda_l k_c - \sum_i \omega_i (\lambda_l k_x x_i + \lambda_l k_y y_i + \lambda_l k_c) \right) \right) &= 0\end{aligned}\quad (2.1.52)$$

Como λ_l , k_x , k_y e k_c são constantes desconhecidas, as multiplicações

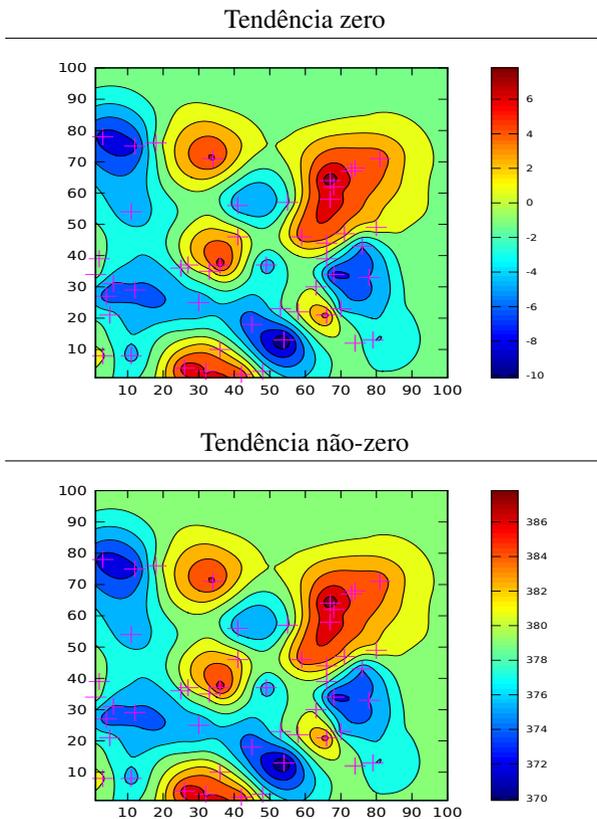


Figura 5: Na Krigagem Ordinária, dados amostrais apenas com médias diferentes podem ser interpolados com resultados semelhantes, apresentando somente a escala diferente, o que não é possível na Krigagem Simples.

$\lambda_i k_x$, $\lambda_i k_y$ e $\lambda_i k_c$ serão referenciadas por λ_x , λ_y e λ_c para simplificar. Então:

$$\nabla \left(\sigma^2 + \left(\lambda_x x_j + \lambda_y y_j + \lambda_c - \sum_i \omega_i (\lambda_x x_i + \lambda_y y_i + \lambda_c) \right) \right) = 0 \quad (2.1.53)$$

A Equação (2.1.29) pode ser escrita como:

$$\begin{aligned} k_x x_j + k_y y_j + k_c - \sum_i \omega_i (k_x x_i + k_y y_i + k_c) &= 0 \\ \sum_i \omega_i (k_x x_i + k_y y_i + k_c) &= k_x x_j + k_y y_j + k_c \\ \sum_i \omega_i k_x x_i + \sum_i \omega_i k_y y_i + \sum_i \omega_i k_c &= k_x x_j + k_y y_j + k_c, \end{aligned} \quad (2.1.54)$$

que é equivalente às três equações:

$$\sum_i \omega_i k_c = k_c \quad (2.1.55)$$

$$\sum_i \omega_i k_x x_i = k_x x_j \quad (2.1.56)$$

$$\sum_i \omega_i k_y y_i = k_y y_j \quad (2.1.57)$$

Como os parâmetros k_c , k_x e k_y podem ser eliminados das equações algebricamente, o resultado é:

$$\sum_i \omega_i = 1 \quad (2.1.58)$$

$$\sum_i \omega_i x_i = x_j \quad (2.1.59)$$

$$\sum_i \omega_i y_i = y_j \quad (2.1.60)$$

Juntam-se as equações (2.1.53), (2.1.58), (2.1.59) e (2.1.60) para for-

mar:

$$\nabla \left(\sigma^2 + \left(\lambda_x x_j + \lambda_y y_j + \lambda_c - \sum_i \omega_i (\lambda_x x_i + \lambda_y y_i + \lambda_c) \right) \right) = 0$$

$$\sum_i \omega_i = 1 \quad (2.1.61)$$

$$\sum_i \omega_i x_i = x_j$$

$$\sum_i \omega_i y_i = y_j$$

Agora, substitui-se σ^2 de (2.1.20) e aplica-se a derivada conforme o Multiplicador de Lagrange, que gera as seguintes equações:

$$\frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2 \gamma_{1j} - \sum_i \omega_i (\lambda_x x_i + \lambda_y y_i + \lambda_c)}{\partial \omega_1} = 0$$

$$\frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2 \gamma_{2j} - \sum_i \omega_i (\lambda_x x_i + \lambda_y y_i + \lambda_c)}{\partial \omega_2} = 0$$

$$\vdots$$

$$\frac{\partial - \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2 \gamma_{nj} - \sum_i \omega_i (\lambda_x x_i + \lambda_y y_i + \lambda_c)}{\partial \omega_n} = 0 \quad (2.1.62)$$

$$\sum_i \omega_i = 1$$

$$\sum_i \omega_i x_i = x_j$$

$$\sum_i \omega_i y_i = y_j,$$

onde n é o número de dados amostrais. Os termos constantes serão zerados pela derivada e não foram mostrados para diminuir o tamanho das equações.

O número 2 acaba cancelado após derivar, além de vários termos que

acabam zerados, resultando nas equações:

$$\begin{aligned}
 \sum_i \omega_i \gamma_{1i} + \lambda_c + \lambda_x x_1 + \lambda_y y_1 &= \gamma_{1j} \\
 \sum_i \omega_i \gamma_{2i} + \lambda_c + \lambda_x x_2 + \lambda_y y_2 &= \gamma_{2j} \\
 &\vdots \\
 \sum_i \omega_i \gamma_{ni} + \lambda_c + \lambda_x x_n + \lambda_y y_n &= \gamma_{nj}
 \end{aligned} \tag{2.1.63}$$

$$\begin{aligned}
 \sum_i \omega_i &= 1 \\
 \sum_i \omega_i x_i &= x_j \\
 \sum_i \omega_i y_i &= y_j,
 \end{aligned}$$

que é equivalente à forma matricial:

$$\overbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} & 1 & x_1 & y_1 \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} & 1 & x_2 & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} & 1 & x_n & y_n \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \dots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \dots & y_n & 0 & 0 & 0 \end{bmatrix}}^{\mathbf{\Gamma}} \overbrace{\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \\ \lambda_c \\ \lambda_x \\ \lambda_y \end{bmatrix}}^{\mathbf{\Omega}_k} = \overbrace{\begin{bmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{nj} \\ 1 \\ x_j \\ y_j \end{bmatrix}}^{\mathbf{\Upsilon}} \tag{2.1.64}$$

Os símbolos $\mathbf{\Gamma}$, $\mathbf{\Omega}_k$ e $\mathbf{\Upsilon}$ estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz $\mathbf{\Gamma}$ e multiplicá-la pela matriz $\mathbf{\Upsilon}$:

$$\begin{aligned}
 \mathbf{\Gamma} \mathbf{\Omega}_k &= \mathbf{\Upsilon} \\
 \mathbf{\Omega}_k &= \mathbf{\Gamma}^{-1} \mathbf{\Upsilon}
 \end{aligned} \tag{2.1.65}$$

A Krigagem Universal é uma boa escolha quando existir realmente uma tendência relacionada a uma direção. Caso contrário, gera resultados piores que a Krigagem Ordinária porque será necessário estimar os parâmetros λ_x e λ_y desnecessariamente, em detrimento de uma melhor estimativa dos pesos ω .

Além disso, esse tipo de Krigagem não está restrita apenas ao polinô-

mio apresentado e qualquer polinômio pode ser usado. É importante que o polinômio utilize informações dependentes de cada ponto para que a tendência fique o mais fiel possível, mas não é um processo obrigatório. Por isso, a Krigagem Simples e a Krigagem Ordinária são apenas casos particulares da Krigagem Universal.

2.1.4 Algoritmo da Krigagem

Os pseudocódigos da Krigagem estão descritos nos algoritmos 1, 2 e 3.

Dados:

Dados amostrais (n dados), representados pela matriz Z . O símbolo $Z[1]$ representa o valor do dado amostral no ponto 1.

Modelo de semivariograma γ já parametrizado.

Funções de Distância, MultiplicaMatrizes, MatrizInversa e MatrizTransposta.

Resultado:

Dados interpolados representados por R . O símbolo $R[j]$ representa o valor interpolado no ponto j .

início

para $a \leftarrow 1$ até n **faça**

para $b \leftarrow 1$ até n **faça**

$\Gamma[a, b] \leftarrow \gamma(\text{Distância}(a, b));$

fim

fim

para cada ponto j dos pontos a serem interpolados **faça**

para $a \leftarrow 1$ até n **faça**

$\Upsilon[a] \leftarrow \gamma(\text{Distância}(a, j));$

fim

$\Omega_k \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Gamma), \Upsilon);$

$R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_k), Z);$

fim

fim

Algoritmo 1: Krigagem Simples

Dados:

Dados amostrais (n dados), representados pela matriz Z . O símbolo $Z[1]$ representa o valor do dado amostral no ponto 1.

Modelo de semivariograma γ já parametrizado.

Funções de Distância, MultiplicaMatrizes, MatrizInversa e MatrizTransposta.

Resultado:

Dados interpolados representados por R . O símbolo $R[j]$ representa o valor interpolado no ponto j .

início

para $a \leftarrow 1$ **até** n **faça**

para $b \leftarrow 1$ **até** n **faça**

$\Gamma[a, b] \leftarrow \gamma(\text{Distância}(a, b));$

fim

$\Gamma[a, n + 1] \leftarrow 1;$

$\Gamma[n + 1, a] \leftarrow 1;$

fim

$\Gamma[n + 1, n + 1] \leftarrow 0;$

para cada ponto j *dos pontos a serem interpolados* **faça**

para $a \leftarrow 1$ **até** n **faça**

$\Upsilon[a] \leftarrow \gamma(\text{Distância}(a, j));$

fim

$\Upsilon[n + 1] \leftarrow 1;$

$\Omega_k \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Gamma), \Upsilon);$

$Z[n + 1] \leftarrow 0;$

$R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_k), Z);$

fim

fim

Algoritmo 2: Krigagem Ordinária

Dados:

Idem ao da Krigagem Simples, com a adição das coordenadas dos pontos, representados pelas letras X e Y . O símbolo $X[a]$ representa a coordenada horizontal do ponto a .

Resultado:

Idem ao da Krigagem Simples.

início

para $a \leftarrow 1$ **até** n **faça**

para $b \leftarrow 1$ **até** n **faça**

$\Gamma[a, b] \leftarrow \gamma(\text{Distância}(a, b));$

fim

$\Gamma[a, n+1] \leftarrow 1;$

$\Gamma[a, n+2] \leftarrow X[a];$

$\Gamma[a, n+3] \leftarrow Y[a];$

$\Gamma[n+1, a] \leftarrow 1;$

$\Gamma[n+2, a] \leftarrow X[a];$

$\Gamma[n+3, a] \leftarrow Y[a];$

fim

para $a \leftarrow n+1$ **até** $n+3$ **faça**

para $b \leftarrow n+1$ **até** $n+3$ **faça**

$\Gamma[a, b] \leftarrow 0;$

fim

fim

para cada *ponto* j *dos pontos* a *serem interpolados* **faça**

para $a \leftarrow 1$ **até** n **faça**

$\Upsilon[a] \leftarrow \gamma(\text{Distância}(a, j));$

fim

$\Upsilon[n+1] \leftarrow 1;$

$\Upsilon[n+2] \leftarrow X[j];$

$\Upsilon[n+3] \leftarrow Y[j];$

$\Omega_k \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Gamma), \Upsilon);$

para $a \leftarrow n+1$ **até** $n+3$ **faça**

$Z[a] \leftarrow 0;$

fim

$R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_k), Z);$

fim

fim

Algoritmo 3: Krigagem Universal

2.1.5 Variância da Krigagem

Uma informação muito útil para uma interpolação é saber qual o intervalo de confiança da interpolação, ou seja, em qual intervalo o valor real provavelmente está. Pela estatística, assumindo um erro gaussiano, há aproximadamente 95% de chance de o valor real estar entre dois desvios padrão abaixo da média e dois desvios padrão acima. O desvio padrão é a raiz da variância. A variância é definida por:

$$\text{var}(z) = E(z - \mu)^2, \quad (2.1.66)$$

ou seja, o valor esperado do quadrado da diferença entre uma variável aleatória z e seu valor esperado μ .

Na Krigagem, espera-se que a média do valor interpolado seja igual ao valor real, já que o valor esperado do erro é zero. Portanto, a Equação (2.1.6), que descreve a função de erro da Krigagem, é, na realidade, a variância do valor interpolado.

Por esse motivo, o símbolo que foi usado para descrevê-la é intencionalmente igual ao símbolo usado normalmente para descrever a variância: σ^2 . Se for possível reescrevê-la em função de variáveis conhecidas, é possível calcular a variância da interpolação e, conseqüentemente, o intervalo de confiança.

Estatisticamente, o valor verdadeiro da interpolação tem aproximadamente 95% de chance de estar entre $z - 2\sqrt{\sigma^2}$ e $z + 2\sqrt{\sigma^2}$, onde z é o valor interpolado. Esse intervalo também é chamado de intervalo de predição (CRESSIE, 1993). O cálculo do σ^2 é diferente em cada tipo de Krigagem.

2.1.5.1 Variância na Krigagem Simples

Aqui é feito o cálculo da variância para o caso mais simples: a Krigagem Simples. As equações (2.1.36) e (2.1.20) são reescritas aqui por conveniência:

$$\sigma^2 = E \left(- \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i^2 \gamma_{ij} \right) \quad (2.1.20)$$

$$\begin{aligned}
\sum_i \omega_i \gamma_{1i} &= \gamma_{1j} \\
\sum_i \omega_i \gamma_{2i} &= \gamma_{2j} \quad (2.1.36) \\
&\vdots \\
\sum_i \omega_i \gamma_{ni} &= \gamma_{nj}
\end{aligned}$$

Agora as variáveis isoladas no lado esquerdo das equações (2.1.36) são substituídas na Equação (2.1.20):

$$\begin{aligned}
\sigma^2 &= E \left(-\sum_i \omega_i \gamma_{ij} + \sum_i \omega_i 2\gamma_{ij} \right) \\
&= E \left(\sum_i \omega_i (-\gamma_{ij}) + \sum_i \omega_i 2\gamma_{ij} \right) \\
&= E \left(\sum_i \omega_i (-\gamma_{ij} + 2\gamma_{ij}) \right) \\
&= E \left(\sum_i \omega_i \gamma_{ij} \right) \quad (2.1.67)
\end{aligned}$$

A Equação (2.1.67) descreve a variância somente em função dos parâmetros conhecidos e do semivariograma, o que torna possível calcular a incerteza.

2.1.5.2 Variância na Krigagem Ordinária

Aqui é feito o cálculo da variância para a Krigagem Ordinária. O mesmo processo descrito na seção anterior também pode ser utilizado para o cálculo da variância para a Krigagem Ordinária. A Equação (2.1.48) é

reescrita aqui por conveniência:

$$\begin{aligned}
 \sum_i \omega_i \gamma_{1i} + \lambda_c &= \gamma_{1j} \\
 \sum_i \omega_i \gamma_{2i} + \lambda_c &= \gamma_{2j} \\
 &\vdots \\
 \sum_i \omega_i \gamma_{ni} + \lambda_c &= \gamma_{nj} \\
 \sum_i \omega_i &= 1,
 \end{aligned} \tag{2.1.48}$$

que podem ser reescritas em:

$$\begin{aligned}
 \sum_i \omega_i \gamma_{1i} &= -\lambda_c + \gamma_{1j} \\
 \sum_i \omega_i \gamma_{2i} &= -\lambda_c + \gamma_{2j} \\
 &\vdots \\
 \sum_i \omega_i \gamma_{ni} &= -\lambda_c + \gamma_{nj} \\
 \sum_i \omega_i &= 1
 \end{aligned} \tag{2.1.68}$$

Agora as variáveis isoladas no lado esquerdo das equações (2.1.68) são substituídas na Equação (2.1.20):

$$\begin{aligned}
 \sigma^2 &= E \left(-\sum_i \omega_i (-\lambda_c + \gamma_{ij}) + \sum_i \omega_i 2\gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c - \gamma_{ij}) + \sum_i \omega_i 2\gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c - \gamma_{ij} + 2\gamma_{ij}) \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c + \gamma_{ij}) \right) \\
 &= E \left(\sum_i \omega_i \lambda_c + \sum_i \omega_i \gamma_{ij} \right).
 \end{aligned} \tag{2.1.69}$$

Como o somatório dos pesos é 1, pela Equação (2.1.68), pode-se simplificar a equação em:

$$\sigma^2 = E \left(\lambda_c + \sum_i \omega_i \gamma_{ij} \right) \quad (2.1.70)$$

A Equação (2.1.70) descreve a variância somente em função dos parâmetros conhecidos e do semivariograma, o que torna possível calcular a incerteza.

2.1.5.3 Variância na Krigagem Universal

Aqui é feito o cálculo da variância para o caso mais complexo: a Krigagem Universal. O mesmo processo descrito na seção anterior também pode ser utilizado para o cálculo da variância para a Krigagem Universal. A Equação (2.1.63) é aqui reescrita por conveniência:

$$\begin{aligned} \sum_i \omega_i \gamma_{1i} + \lambda_c + \lambda_x x_1 + \lambda_y y_1 &= \gamma_{1j} \\ \sum_i \omega_i \gamma_{2i} + \lambda_c + \lambda_x x_2 + \lambda_y y_2 &= \gamma_{2j} \\ &\vdots \\ \sum_i \omega_i \gamma_{ni} + \lambda_c + \lambda_x x_n + \lambda_y y_n &= \gamma_{nj} \end{aligned} \quad (2.1.63)$$

$$\begin{aligned} \sum_i \omega_i &= 1 \\ \sum_i \omega_i x_i &= x_j \\ \sum_i \omega_i y_i &= y_j, \end{aligned}$$

que podem ser reescritas em:

$$\begin{aligned}
 \sum_i \omega_i \gamma_{1i} &= -(\lambda_c + \lambda_x x_1 + \lambda_y y_1) + \gamma_{1j} \\
 \sum_i \omega_i \gamma_{2i} &= -(\lambda_c + \lambda_x x_2 + \lambda_y y_2) + \gamma_{2j} \\
 &\vdots \\
 \sum_i \omega_i \gamma_{ni} &= -(\lambda_c + \lambda_x x_n + \lambda_y y_n) + \gamma_{nj} \\
 \sum_i \omega_i &= 1 \\
 \sum_i \omega_i x_i &= x_j \\
 \sum_i \omega_i y_i &= y_j,
 \end{aligned} \tag{2.1.71}$$

Agora variáveis isoladas no lado esquerdo das equações (2.1.71) são substituídas na Equação (2.1.20):

$$\begin{aligned}
 \sigma^2 &= E \left(-\sum_i \omega_i (-(\lambda_c + \lambda_x x_i + \lambda_y y_i) + \gamma_{ij}) + \sum_i \omega_i 2\gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c + \lambda_x x_i + \lambda_y y_i - \gamma_{ij}) + \sum_i \omega_i 2\gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c + \lambda_x x_i + \lambda_y y_i - \gamma_{ij} + 2\gamma_{ij}) \right) \\
 &= E \left(\sum_i \omega_i (\lambda_c + \lambda_x x_i + \lambda_y y_i + \gamma_{ij}) \right) \\
 &= E \left(\sum_i \omega_i \lambda_c + \sum_i \omega_i \lambda_x x_i + \sum_i \omega_i \lambda_y y_i + \sum_i \omega_i \gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i \lambda_c + \lambda_x \sum_i \omega_i x_i + \lambda_y \sum_i \omega_i y_i + \sum_i \omega_i \gamma_{ij} \right) \\
 &= E \left(\sum_i \omega_i \lambda_c + \lambda_x x_j + \lambda_y y_j + \sum_i \omega_i \gamma_{ij} \right).
 \end{aligned} \tag{2.1.72}$$

Como o somatório dos pesos é 1, pela Equação (2.1.71), pode-se sim-

plificar a equação em:

$$\sigma^2 = E \left(\lambda_c + \lambda_x x_j + \lambda_y y_j + \sum_i \omega_i \gamma_i \right) \quad (2.1.73)$$

A Equação (2.1.73) descreve a variância somente em função dos parâmetros conhecidos e do semivariograma, o que torna possível calcular a incerteza.

2.1.6 Semivariograma

Nas equações desenvolvidas da Krigagem foi utilizado o valor do semivariograma. Como apresentado anteriormente, o semivariograma é definido por (CRESSIE, 1993):

$$\gamma(h) = \frac{E(z_a - z_b)^2}{2}, \quad (2.1.74)$$

que é a metade da diferença quadrática esperada entre dois pontos, representados por a e b , cuja distância entre si é h . O semivariograma descreve o grau de dependência espacial de um processo estocástico como a interpolação.

A Krigagem assume que valores próximos são semelhantes, o que implica que o semivariograma deve retornar valores baixos. Para distâncias maiores, espera-se que o semivariograma retorne valores mais altos, já que o grau de dependência dos dados é menor (CRESSIE, 1993; GNEITING; SAS-VÁRI; SCHLATHER, 2000; GRINGARTEN; DEUTSCH, 2001). A função acima precisa ser modelada pois, com os dados amostrais, tem-se somente essa função definida para algumas distâncias.

2.1.6.1 Semivariograma experimental

O modelo a ser criado para o semivariograma precisa ter seus parâmetros ajustados, já que é difícil conhecer esses parâmetros antecipadamente. Os dados amostrais são usados para este ajuste (CRESSIE, 1993). Com esses dados é gerado o semivariograma experimental, também chamado de semivariograma empírico, onde todas os pontos amostrais a e b com a mesma distância h são utilizadas para calcular a média da diferença quadrática entre seus valores z_a e z_b . Haverá uma quantidade finita de distâncias no semivariograma experimental e, por isso, ele será usado unicamente para ajustar o modelo de semivariograma. O semivariograma experimental é representado

por $\gamma^{exp}(h)$ neste trabalho. A Figura 6 mostra um exemplo de um semivariograma experimental em vermelho e um possível modelo de semivariograma ajustado a ele em azul.

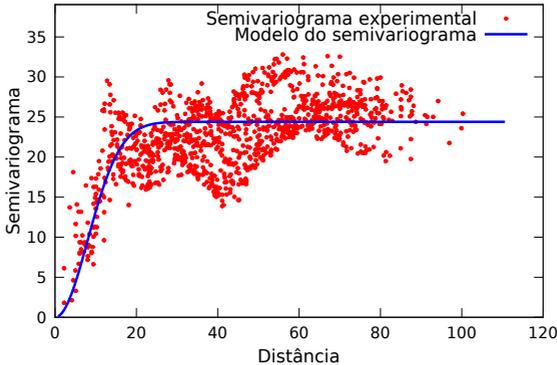


Figura 6: Exemplo de semivariograma experimental

As técnicas para ajustar o modelo de semivariograma com base no semivariograma experimental serão mostradas na Seção 2.1.6.4.

2.1.6.2 Modelo de semivariograma

Existem diversos modelos possíveis de serem usados para representar o semivariograma, também chamados de semivariograma teórico. Normalmente estes modelos possuem parâmetros com significados, conforme mostrado na Figura 7. O *alcance* indica a distância em que o modelo passa a responder um valor aproximadamente constante. Isso não se aplica em modelos que crescem indefinidamente, como o modelo Linear da Equação (2.1.75). O *limiar* é a saída do modelo quando a distância é maior que o *alcance*. A *pepita* é o valor do modelo quando a distância é próxima de zero (CRESSIE, 1993; DALEZIOS; LOUKAS; BAMPZELIS, 2002).

Alguns deles são mostrados abaixo, e as versões gráficas deles são mostradas pela Figura 8.

- **Linear:**

$$\gamma(h) = \begin{cases} 0 & \text{se } h = 0, \\ pepita + \frac{limiar-pepita}{alcance}h & \text{se } h \neq 0. \end{cases} \quad (2.1.75)$$

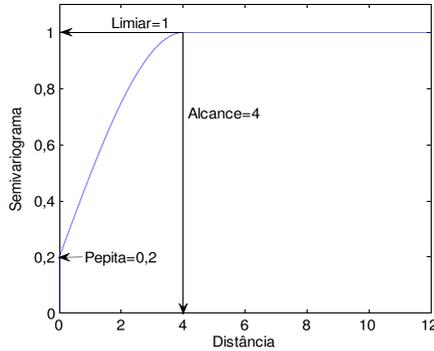


Figura 7: Parâmetros do modelo de semivariograma.

- **Esférico:**

$$\gamma(h) = \begin{cases} 0 & \text{se } h = 0, \\ \text{pepita} + \frac{\text{limiar} - \text{pepita}}{2\text{alcance}} \left(3h - \frac{h^3}{\text{alcance}^2} \right) & \text{se } 0 < h < \text{alcance}, \\ \text{limiar} & \text{se } h \geq \text{alcance}. \end{cases} \quad (2.1.76)$$

- **Gaussiano:**

$$\gamma(h) = \begin{cases} 0 & \text{se } h = 0, \\ \text{pepita} + (\text{limiar} - \text{pepita}) \left(1 - e^{-\frac{h^2}{\text{alcance}}} \right) & \text{se } h \neq 0. \end{cases} \quad (2.1.77)$$

O parâmetro h é a distância. Os parâmetros $pepita$, $limiar$ e $alcance$ precisam ser estimados de forma que o modelo gerado se aproxime do semivariograma experimental.

O efeito de cada um destes parâmetro na Krigagem está demonstrado na Figura 9. Nessa figura, um parâmetro é variado para verificar o seu efeito e todos os outros parâmetros ficam fixos. É visível que o parâmetro $limiar$ e $pepita$ não influenciam a interpolação, tendo como resultado valores idênticos em todos os pontos interpolados. Isso ocorre porque o efeito deles na matriz Γ e Υ da Krigagem serão proporcionais, e os pesos irão ajustar o mesmo resultado. Já o parâmetro $alcance$ acaba determinando quais dados amostrais têm mais influência no resultado final, porque todos os dados amostrais além do $alcance$ têm o mesmo valor de semivariograma e, portanto, sem influência no resultado. Isso faz com que pontos interpolados cujo $alcance$

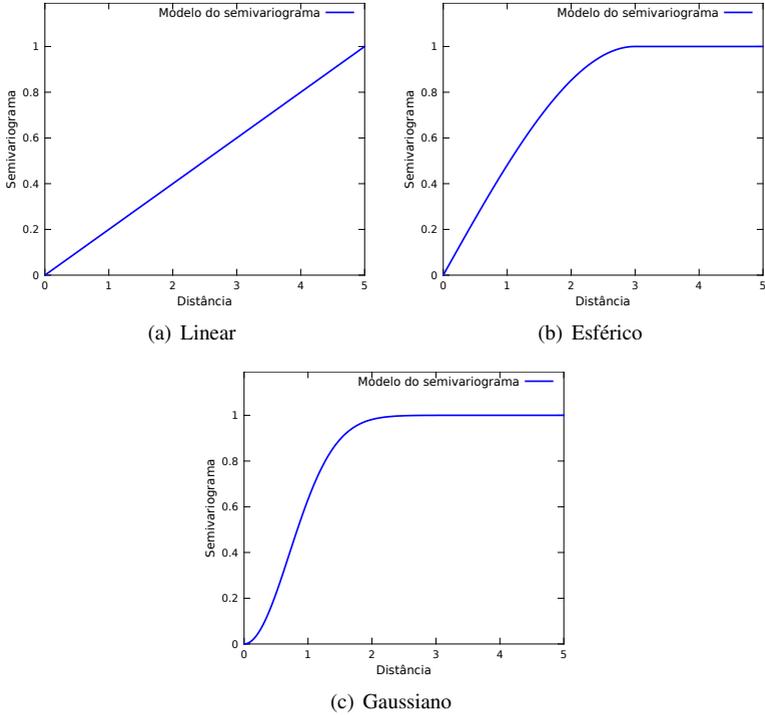


Figura 8: Alguns exemplos gráficos de modelos de semivariogramas.

não chega a nenhum ponto amostral tenham a mesma interpolação, gerando uma grande superfície verde representando o mesmo valor no exemplo onde o *alcance* = 10. Apesar de não ter efeito sobre a interpolação, os parâmetros *limiar* e *pepita* são utilizados para que o semivariograma tenha um modelo ajustado que represente bem o semivariograma experimental.

2.1.6.3 Validade do semivariograma

Para que um modelo de semivariograma possa ser utilizado pela Krigagem, ele precisa obedecer a algumas restrições (CRESSIE, 1993):

- nunca pode retornar um valor negativo, já que a própria definição de semivariograma é um valor ao quadrado;
- deve retornar zero quando é calculado o semivariograma da distância

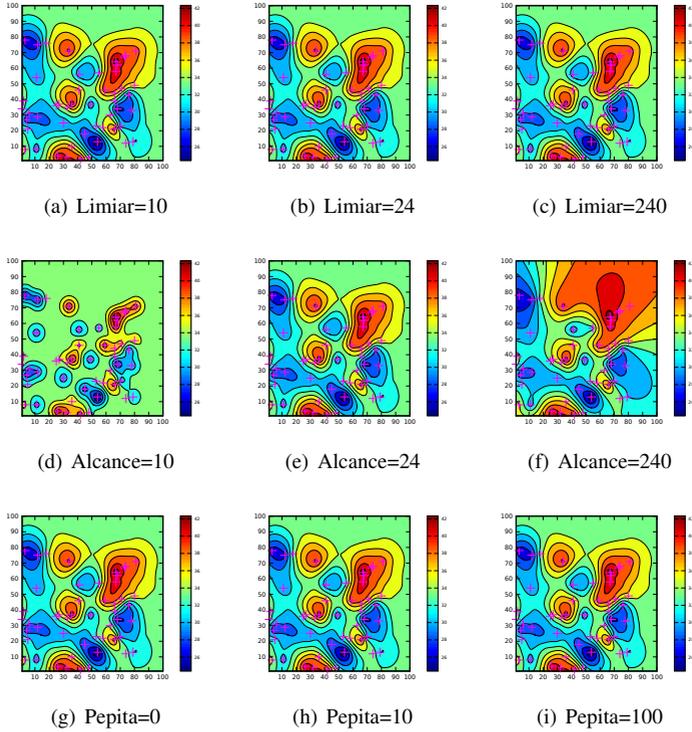


Figura 9: Comportamento da Krigagem variando seus parâmetros.

zero, já que pela própria definição do semivariograma, quando a distância é zero, os valores z serão iguais e a subtração resultará sempre em zero;

- deve ser *condicionalmente negativo definido*, o que significa que ele respeita a seguinte restrição:

$$\sum_{i1} \sum_{i2} a_{i1} a_{i2} 2\gamma_{f1,i2} \leq 0$$

$$\sum_i a_i = 0 \quad (2.1.78)$$

onde $i1$ e $i2$ são dois pontos quaisquer, γ é o modelo de semivariograma, aplicado aos pontos $i1$ e $i2$, e a é um número real qualquer, relacionado aos

pontos $i1$ e $i2$.

Para esclarecer mais essa restrição, basta rever a fórmula da variância que se minimiza na Krigagem, da Equação (2.1.20), reescrita aqui por conveniência:

$$\sigma^2 = E \left(- \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2\gamma_{ij} \right) \quad (2.1.20)$$

A equação é manipulada algebricamente para que chegue a um formato em que possa ser mais fácil verificar a restrição:

$$\begin{aligned} \sigma^2 &= E \left(- \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} \gamma_{i1,i2} + \sum_i \omega_i 2\gamma_{ij} \right) \\ &= E \left(- \sum_{i1} \omega_{i1} \left(\sum_{i2} \omega_{i2} \gamma_{i1,i2} - 2\gamma_{i1,j} \right) \right) \end{aligned} \quad (2.1.79)$$

Como a soma dos pesos é igual a 1, conforme visto na Seção 2.1, é possível fazer:

$$\begin{aligned} \sigma^2 &= E \left(- \sum_{i1} \omega_{i1} \left(\sum_{i2} \omega_{i2} \gamma_{i1,i2} - 2\gamma_{i1,j} \sum_i \omega_i \right) \right) \\ &= E \left(- \sum_{i1} \omega_{i1} \sum_{i2} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) \right) \\ &= E \left(- \sum_{i1} \sum_{i2} \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) \right), \end{aligned} \quad (2.1.80)$$

onde se pode afirmar, pela definição de variância, que:

$$\begin{aligned} - \sum_{i1} \sum_{i2} \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) &\geq 0 \\ \sum_{i1} \sum_{i2} \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) &\leq 0 \end{aligned} \quad (2.1.81)$$

A Equação (2.1.81) é quase idêntica à restrição de ser *condicionalmente negativo definido*, onde os pesos ω fazem o papel dos números reais a , com a diferença de que soma dos pesos ω é igual a 1 e a soma dos números a é igual a zero, além do do termo $2\gamma_{i1,i2}$ ter sido substituído por $\gamma_{i1,i2} - 2\gamma_{i1,j}$.

O termo “variograma” é definido pelo dobro do semivariograma. Como a desigualdade exige que o valor seja negativo, subtrair um variograma, que

por definição é sempre positivo, só vai garantir ainda mais que a desigualdade seja aceita, já que quando $i1 = i2$, a multiplicação $\omega_{i1} \omega_{i2} \geq 0$ e $\gamma_{i1,i2} = 0$, o que deixará o somatório ainda mais negativo quando for feita a multiplicação pelo termo $-2\gamma_{i1,j}$. A multiplicação por 2 feita na restrição não altera a igualdade e pode ser ignorada por ser apenas a multiplicação de um valor positivo.

Agora, é feito um ajuste nos pesos ω para que soma deles se torne zero, subtraindo uma fração de 1 igual em cada multiplicação entre os pesos, onde a soma de todas essas frações será 1:

$$\begin{aligned} \sum_{i1} \sum_{i2} \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) &\leq 0 \\ \sum_{i1} \sum_{i2} \left(1 - \frac{1}{n^2}\right) \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) &\leq 0 \\ \left(1 - \frac{1}{n^2}\right) \sum_{i1} \sum_{i2} \omega_{i1} \omega_{i2} (\gamma_{i1,i2} - 2\gamma_{i1,j}) &\leq 0 \end{aligned} \quad (2.1.82)$$

Como $1 - \frac{1}{n^2}$ é sempre um valor entre 0 e 1, essa multiplicação não altera a desigualdade, onde se conclui que um semivariograma que respeite a restrição de ser *condicionalmente negativo definido* garante que a variância gerada por ele seja sempre positiva.

Um modelo de semivariograma contínuo, não precisando ser contínuo ao redor de zero, e que satisfaça a restrição $\gamma(0) = 0$, já é um semivariograma válido e *condicionalmente negativo definido*, segundo Cressie (CRESSIE, 1993).

2.1.6.4 Ajuste do semivariograma

O ajuste do modelo pode ser feito de forma iterativa baseada em uma função de erro que utiliza os dados do semivariograma experimental para ser computada. Existem diversas funções de erro que pode ser utilizada para esse ajuste. Alguns exemplos de função de erro para ajuste do semivariograma são (CRESSIE, 1993):

- Erro quadrático: esse é o método tradicional, que calcula o quadrado do erro. Um inconveniente deste método é que ele atribui muita importância quando o erro é grande:

$$\text{erro} = E \left(\gamma(h) - \overset{\text{exp}}{\gamma}(h) \right)^2 \quad (2.1.83)$$

- Erro relativo: esse método dá importância igual às pequenas e grandes distâncias, pois é medida a relação entre o semivariograma experimental e o teórico. Um pequeno erro em pequena distância pode gerar um grande erro relativo:

$$\text{erro} = E \left(\frac{\gamma(h)}{\gamma^{exp}(h)} - 1 \right)^2 \quad (2.1.84)$$

- Outro método, descrito na equação a seguir, tem como meta diminuir o efeito dos valores discrepantes. A letra n representa o número de pares de pontos da amostra com a mesma distância h :

$$\text{erro} = \frac{1}{0,457 + \frac{0,494}{n}} \left(\frac{1}{n} \sum_h |\gamma(h) - \gamma^{exp}(h)|^{1/2} \right)^4 \quad (2.1.85)$$

Os parâmetros do semivariograma que minimizam a função de erro precisam ser estimados. Isso pode ser feito manualmente ou através de um método iterativo, como Levenberg-Marquadt (MARQUARDT, 1963; PRESS et al., 1992), Região de Confiança, mais conhecido como *Trust-Region*, e Têmpera Simulada, mais conhecido como *Simulated Annealing* (KIRKPATRICK; GELATT; VECCHI, 1983).

2.1.6.5 Anisotropia em semivariograma

Em alguns conjuntos de dados observa-se que a distância em uma determinada direção tem menor ou maior influência no semivariograma. A Figura 10 mostra um exemplo deste tipo de dado. Cada ponto representa um dado localizado espacialmente. A cor representa o valor do dado. Horizontalmente o valor varia muito pouco enquanto verticalmente varia bastante.

Nesse caso é necessário estimar mais dois parâmetros no modelo de semivariograma para obter mais precisão: o ângulo e o fator de anisotropia (CRESSIE, 1993). Esses dois parâmetros entram no cálculo da distância h dos modelos de semivariograma descritos na Seção 2.1.6.2.

Uma função de distância que não considera o ângulo, dita isotrópica,

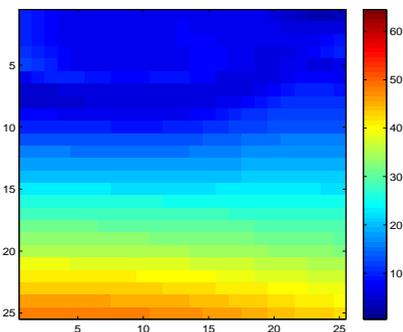


Figura 10: Exemplo de dados anisotrópicos.

é a Distância Euclidiana:

$$\begin{aligned}
 h_x &= x_a - x_b \\
 h_y &= y_b - y_a \\
 \text{distância}(a, b) &= \sqrt{h_x^2 + h_y^2},
 \end{aligned} \tag{2.1.86}$$

onde a e b são dois pontos diferentes, x e y são respectivamente as coordenadas horizontal e vertical.

Uma função de distância anisotrópica que pode ser usada para o caso específico da Figura 10:

$$\begin{aligned}
 h_x &= \frac{1}{r}(x_a - x_b) \\
 h_y &= y_a - y_b \\
 \text{distância}(a, b) &= \sqrt{h_x^2 + h_y^2},
 \end{aligned} \tag{2.1.87}$$

onde r é o fator de anisotropia e as demais variáveis tem o mesmo significado da equação anterior. Quanto maior o r , menor será o h_x , conseqüentemente menor a distância e, por fim, menor o resultado do modelo semivariograma, que normalmente é proporcional à distância. Enquanto a variação na distância vertical continuará com o efeito normal sobre o semivariograma.

Para generalizar um pouco mais a Equação (2.1.87) pode-se adicionar

o parâmetro de ângulo, representado por α :

$$\begin{aligned} h_x &= \frac{1}{r} ((x_a - x_b) \cos \alpha - (y_a - y_b) \sin \alpha) \\ h_y &= (y_a - y_b) \cos \alpha + (x_a - x_b) \sin \alpha \quad (2.1.88) \\ \text{distância}(a, b) &= \sqrt{h_x^2 + h_y^2} \end{aligned}$$

No ângulo α o semivariograma irá variar menos, mas essa variação irá subir gradualmente até chegar ao ângulo perpendicular a α . Quando o ângulo α for zero, haverá menor variação horizontal, como na Equação (2.1.87). Quando o ângulo α for $\pi/2$, haverá menor variação verticalmente. O fator de anisotropia r deve ser maior que 1 para obter o efeito de anisotropia descrito. Quando o r for igual a 1, a distância fica idêntica a Distância Euclidiana, descrita na Equação (2.1.86).

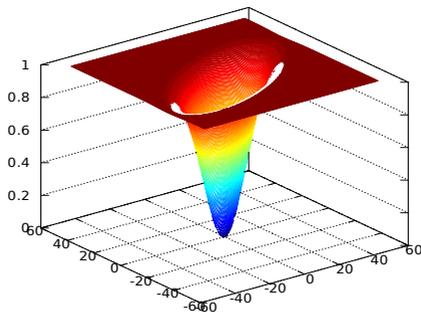
Um exemplo comparativo de um semivariograma isotrópico e anisotrópico pode ser visto na Figura 11, onde os eixos horizontais e verticais representam a distância. Quanto mais afastado do centro, maior a distância. O eixo de altura do gráfico representa o valor do semivariograma. Em um modelo isotrópico, o gráfico tem formato de um círculo, enquanto em um modelo anisotrópico, o gráfico tem formato de uma elipse. É possível perceber que o ângulo formado pelas distâncias nos eixos horizontais e verticais não influenciam o valor do semivariograma isotrópico, diferentemente do anisotrópico.

A Figura 12 mostra o que representam os parâmetros do semivariograma anisotrópico. Neste gráfico, o semivariograma é visto de cima, e as cores representam os seus valores. No centro a distância também é zero. No ângulo, formado a partir do eixo horizontal no sentido anti-horário, indicado por α , o semivariograma varia mais lentamente. Não é possível visualizar o parâmetro *pepita*, que é uma constante somada em toda superfície do gráfico com exceção da distância igual a zero. O parâmetro *alcance* é medido perpendicularmente ao ângulo α e na direção do ângulo α o semivariograma cresce até a distância formada pela multiplicação dos parâmetros *alcance* com r . O parâmetro *limiar* é o valor mais alto do semivariograma, representado pela cor vermelho-escuro.

2.1.7 Covariograma

Outro tipo de modelo que pode ser utilizado na Krigagem é o covariograma. Diferente do semivariograma, ele tenta modelar a correlação entre os valores em função da distância. Quando a distância entre dois pontos é zero, o covariograma responde o valor máximo, já que o valor de um ponto

Semivariograma anisotrópico



Semivariograma Isotrópico

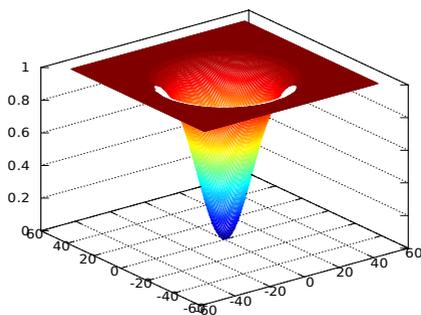


Figura 11: Exemplo de um modelo de semivariograma isotrópico e anisotrópico.

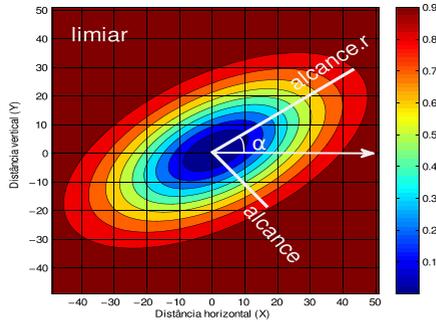


Figura 12: Exemplo de um modelo de semivariograma anisotrópico e seus parâmetros.

tem alta correlação consigo mesmo. À medida que a distância dos pontos aumenta, essa correlação vai diminuindo até chegar a zero, indicando nenhuma correlação entre eles.

O semivariograma e o covariograma possuem a relação (CRESSIE, 1993):

$$\gamma(h) = C(0) - C(h) \quad (2.1.89)$$

onde γ é o semivariograma, C é o covariograma. Porém, essa conversão só pode ser feita quando a tendência é igual para todos os dados, como na Krigagem Simples e na Krigagem Ordinária.

Como o valor de $C(0)$ é o valor máximo da covariância, ela é o *limiar*. Então, a relação entre covariância e semivariograma pode ser reescrita para:

$$C(h) = \text{limiar} - \gamma(h) \quad (2.1.90)$$

A Equação (2.1.37) tem o cálculo dos pesos que minimizam o erro da Krigagem. Como os dois lados da equação possuem apenas semivariogramas, utilizar um covariograma ao invés de um semivariograma é equivalente a multiplicar os dois lados da equação por -1 e somar uma constante *limiar*, o que não altera a igualdade. Por isso, utilizar covariogramas em dados estacionários resulta em pesos iguais e interpolação igual à que utiliza semivariograma.

Apesar da equivalência na interpolação em dados estacionários, onde a tendência é igual para todos os dados, o covariograma ainda é mais limitado que o semivariograma, já que não é possível utilizá-lo quando não existe o

limiar, como é o caso do modelo linear, descrito na Equação (2.1.75). Além disso, a variância da Krigagem, descrita na Seção 2.1.5 sofre alterações, já que ela foi deduzida a partir do semivariograma.

A Figura 13 mostra um exemplo de um covariograma e o modelo de semivariograma equivalente para a Krigagem.

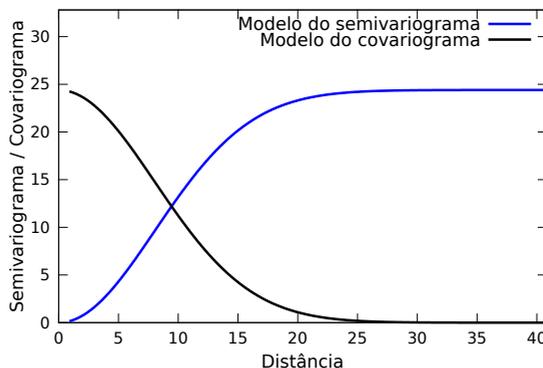


Figura 13: Exemplo de um modelo de covariograma comparado ao seu semivariograma equivalente.

2.2 REDES RBF

As Redes RBF são um tipo de Redes Neurais (BROOMHEAD; LOWE, 1988; HAYKIN, 2008; BISHOP, 1996). Elas utilizam funções de base radial como neurônios, distribuídos espacialmente. Essas funções têm como entrada uma distância. Alguns exemplos de funções de base radial são mostrados abaixo (CARR; FRIGHT; BEATSON, 1997):

$$\text{Linear: } \phi(h) = ah \quad (2.2.1)$$

$$\text{Gaussiana: } \phi(h) = e^{-\frac{h^2}{a}} \quad (2.2.2)$$

$$\text{Multi-quadrática: } \phi(h) = \sqrt{a^2 + (bh)^2}, \quad (2.2.3)$$

onde h é distância. Alguns parâmetros precisam ser estimados em cada função. A função Gaussiana é bastante utilizada e seu parâmetro recebe um nome especial: *spread*. Quanto maior este parâmetro, mais suave é a interpolação.

A Figura 14 mostra a representação gráfica destes modelos.

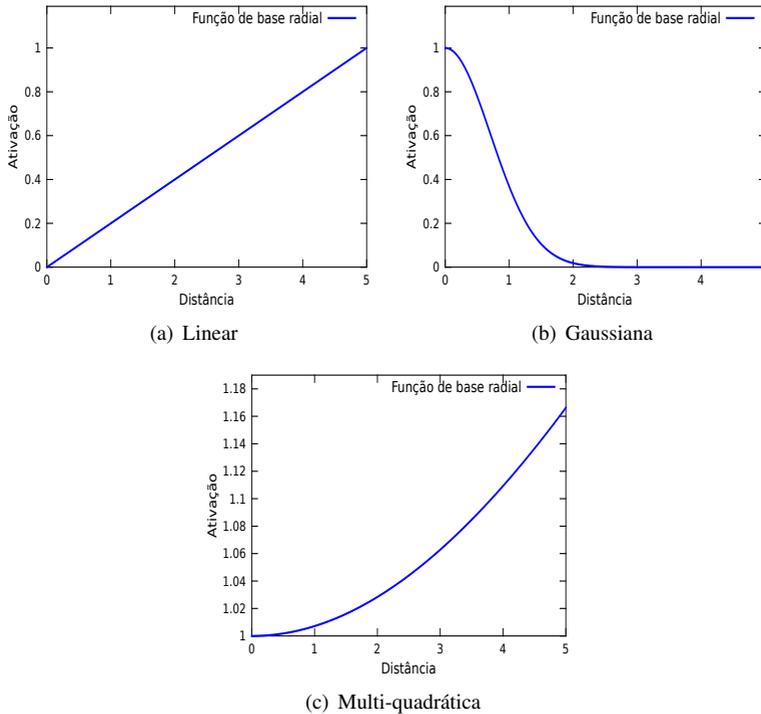


Figura 14: Alguns exemplos gráficos de modelos de semivariogramas.

Na interpolação das Redes RBF, o ponto a ser interpolado pode ativar mais ou menos cada neurônio, dependendo da distância entre o ponto e o centro da função de base radial de cada neurônio, o que resulta em maior ou menor contribuição do neurônio para o resultado final. A ativação é medida pela função de base radial. É mais intuitivo pensar que quanto mais próximo o ponto a ser interpolado está do centro da função de base radial do neurônio, mais o neurônio é ativado, e que conforme ele se distancia do centro de base radial do neurônio, menos ativado ele é, até chegar em zero de ativação. É discutido na Seção 2.2.1 como são definidos os neurônios e seus parâmetros.

A Figura 15 mostra um exemplo de diagrama da arquitetura das Redes RBF. Neste exemplo, existem duas dimensões na entrada: as coordenadas horizontal e vertical, x e y , resumidas no ponto i . É utilizada a distância entre o centro da função de base radial de cada neurônio e a entrada i para calcular a ativação. Depois, cada valor resultante é multiplicado por um peso ω e

somado para gerar a saída z_j . A notação $\phi_{a,b}$ representa $\phi(\text{Distância}(a,b))$.

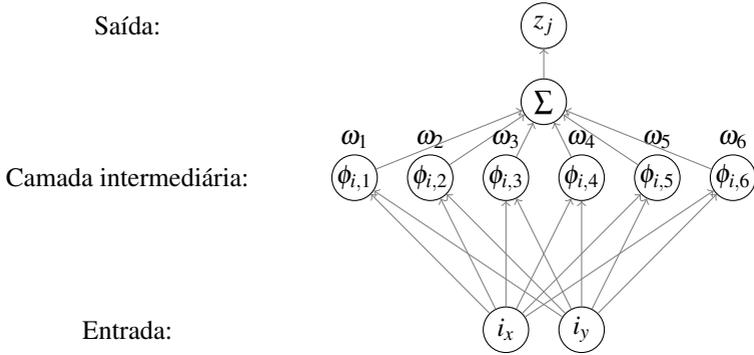


Figura 15: Exemplo de diagrama de uma Rede RBF

A Figura 16 mostra um exemplo gráfico de uma interpolação RBF que utiliza função de base radial gaussiana e somente uma entrada. A posição de cada neurônio é o centro de sua função de base radial. A altura da função de base radial é determinada pelo peso ω , e a abertura pelo parâmetro *spread*. Quanto maior o parâmetro *spread*, maior será a abertura das gaussianas e mais suave será a interpolação final. Quando o parâmetro é pequeno, as gaussianas não se interpoem e o resultado são várias gaussianas curtas espalhadas, e com valores zero onde não há gaussianas perto.

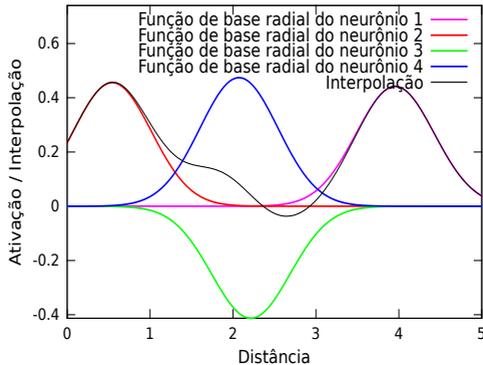


Figura 16: Exemplo da interpolação RBF com somente uma entrada.

A equação de interpolação das Redes RBF é:

$$z_j = \sum_i \omega_i \phi_{ij}, \quad (2.2.4)$$

onde ϕ é a função de base radial, que representa a ativação do neurônio onde $\phi_{ij} = \phi(\text{Distância}(i, j))$, i é o centro da função de base radial de um neurônio, j é o ponto a ser interpolado e ω são os pesos que são calculados para realizar a interpolação. Esse cálculo será descrito na Seção 2.2.1.

2.2.1 Treinamento

O treinamento das Redes RBF define a posição do centro da função de base radial de cada neurônio, a quantidade de neurônios a serem utilizados, os parâmetros da função de base radial (*spread*) e os pesos ω da Equação (2.2.4).

Quanto à definição do centro da função de base radial, as Redes RBF normalmente agrupam os dados amostrais espacialmente próximos e definem uma só função de base radial para representá-los. Esse agrupamento é feito utilizando algoritmos como o K-means (HARTIGAN; WONG, 1979), que cria grupos aleatórios e iterativamente vai separando os elementos através da proximidade espacial. Outra forma é posicionar os centros das RBF aleatoriamente e deixar que o ajuste dos pesos e dos parâmetros sejam suficiente para que a interpolação fique adequada.

Este trabalho analisa as Redes RBF Exatas, que posiciona o centro da função RBF sobre os dados amostrais, formando um neurônio por dado amostral.

Existem diversas formas de ajustar o parâmetro da função de ativação, conhecido como *spread*. Uma forma simples é definir um valor baseado no número de neurônios e na distância média entre os centros das funções de base radial dos neurônios. Podem ser utilizados também métodos de aproximação de parâmetros.

Os pesos podem ser ajustados utilizando algum método iterativo de aproximação de parâmetros. A função de erro que o método terá que minimizar pode ser o erro quadrado entre o resultado obtido da interpolação e o dado real, disponível nos dados amostrais. Alguns exemplos de métodos de aproximação de parâmetros: Levenberg-Marquadt (MARQUARDT, 1963; PRESS et al., 1992), Gradiente-descendente (MORDECAI, 2003) e Têmpera Simulada, mais conhecido como *Simulated Annealing* (KIRKPATRICK; GELATT; VECCHI, 1983).

Quando o centro das funções RBF são posicionados sobre os dados amostrais, o cálculo dos pesos é mais simples, bastando resolver um sistema

de equações lineares, pois há a mesma quantidade de dados amostrais e variáveis a serem estimadas. Cada equação é a interpolação, descrita na Equação (2.2.4), sendo igualada a um dado amostral:

$$\begin{aligned} \sum_i \omega_i \phi_{1i} &= z_1 \\ \sum_i \omega_i \phi_{2i} &= z_2 \\ &\vdots \\ \sum_i \omega_i \phi_{ni} &= z_n, \end{aligned} \quad (2.2.5)$$

que pode ser reescrito em forma matricial:

$$\overbrace{\begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1} & \phi_{n2} & \dots & \phi_{nn} \end{bmatrix}}^{\mathbf{\Phi}} \overbrace{\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix}}^{\mathbf{\Omega}_r} = \overbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}}^{\mathbf{Z}}, \quad (2.2.6)$$

onde ϕ é a função de ativação, ω são os pesos a serem calculados, n é o número de neurônios e $\phi_{ab} = \phi(\text{distancia}(a,b))$, sendo a e b os centros da função de base radial de dois neurônios. Os símbolos $\mathbf{\Phi}$, $\mathbf{\Omega}_r$ e \mathbf{Z} estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz $\mathbf{\Phi}$ e multiplicá-la pela matriz \mathbf{Z} :

$$\begin{aligned} \mathbf{\Phi} \mathbf{\Omega}_r &= \mathbf{Z} \\ \mathbf{\Omega}_r &= \mathbf{\Phi}^{-1} \mathbf{Z} \end{aligned} \quad (2.2.7)$$

2.2.2 Variações

Quando o dado a ser interpolado está longe de todos os neurônios, as Redes RBF respondem com zero, já que nenhum neurônio é ativado. Para evitar esse comportamento, existem variações das Redes RBF. Uma variação adiciona um polinômio à interpolação (CARR; FRIGHT; BEATSON, 1997; SATHYANARAYANAMURTHY; CHINNAM, 2009). Esse polinômio é ajustado para representar a tendência global da interpolação enquanto os neurônios são ajustados para representar as particularidades locais. Esse va-

lor adicionado também pode ser chamado de viés, mais conhecido como *bias*. Um exemplo da arquitetura com o *bias* pode ser visto na Figura 17. Neste exemplo, existem duas dimensões na entrada: as coordenadas horizontal e vertical, x e y , resumidas no ponto i . É utilizada a distância entre o centro da função de base radial de cada neurônio e a entrada i para calcular a ativação. Depois, cada valor resultante é multiplicado por um peso ω e somado ao *bias* para gerar a saída z_j .

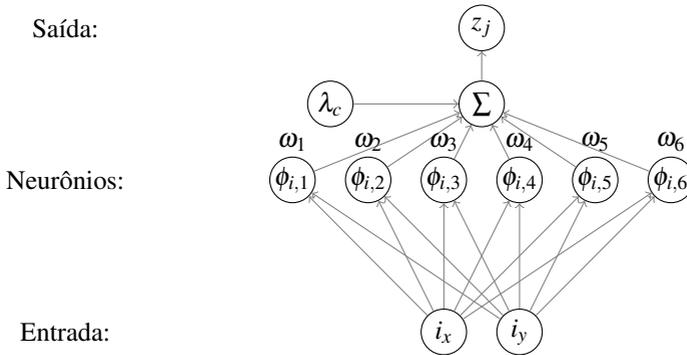


Figura 17: Exemplo de diagrama de uma Rede RBF com *bias*.

Caso esse polinômio seja uma função constante, a Equação (2.2.4), que é a interpolação das Redes RBF, se torna:

$$z_j = \lambda_c + \sum_i \phi_{ij} \omega_i \quad (2.2.8)$$

onde λ_c é o novo parâmetro a ser calculado, que representa a função constante.

Com esse parâmetro extra, a Equação (2.2.6), que calcula os pesos ω , pode ser escrita como:

$$\begin{array}{c} \Phi \\ \left[\begin{array}{cccc|c} \phi_{11} & \phi_{12} & \dots & \phi_{1n} & 1 \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & 1 \\ \phi_{n1} & \phi_{n2} & \dots & \phi_{nn} & 1 \\ \hline 1 & 1 & 1 & 1 & 0 \end{array} \right] \end{array} \begin{array}{c} \Omega_r \\ \left[\begin{array}{c} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \\ \lambda_c \end{array} \right] \end{array} = \begin{array}{c} Z \\ \left[\begin{array}{c} z_1 \\ z_2 \\ \vdots \\ z_n \\ 0 \end{array} \right] \end{array} \quad (2.2.9)$$

Foi adicionada uma linha na matriz, que é transposta à coluna adicionada pelo novo parâmetro. Isso é necessário para que a equação seja resolvida. Os

símbolos Φ , Ω_r e Z estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz Φ e multiplicá-la pela matriz Z :

$$\begin{aligned}\Phi\Omega_r &= Z \\ \Omega_r &= \Phi^{-1}Z\end{aligned}\quad (2.2.10)$$

É possível usar um polinômio com mais termos, que utiliza as coordenadas horizontal e vertical, x e y , de cada neurônio e do ponto a ser interpolado. Dessa forma, a equação de interpolação das Redes RBF pode ser escrita como:

$$z_j = \lambda_c + x_j\lambda_x + y_j\lambda_y + \sum_i \phi_{ij}\omega_i \quad (2.2.11)$$

onde os parâmetros x e y são respectivamente as coordenadas horizontal e vertical do ponto a ser interpolado, λ_x e λ_y são os novos parâmetros a serem calculados. Os demais símbolos são os mesmos da Equação (2.2.8). Assim, o sistema de equação linear fica

$$\begin{array}{c} \overbrace{\begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} & 1 & x_1 & y_1 \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} & 1 & x_2 & y_2 \\ \vdots & \vdots & \ddots & \vdots & 1 & \vdots & \vdots \\ \phi_{n1} & \phi_{n2} & \dots & \phi_{nn} & 1 & x_n & y_n \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \dots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \dots & y_n & 0 & 0 & 0 \end{bmatrix}}^{\Phi} \overbrace{\begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \\ \lambda_c \\ \lambda_x \\ \lambda_y \end{bmatrix}}^{\Omega_r} = \overbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \\ 0 \\ 0 \\ 0 \end{bmatrix}}^Z \end{array} \quad (2.2.12)$$

Novamente foram adicionadas linhas extras na matriz, que é a transposta das colunas extras adicionadas pelos novos parâmetros. Isso é exigido para que a equação seja resolvida. Os símbolos Φ , Ω_r e Z estão representando os elementos da equação de forma simplificada. Para encontrar os pesos, basta inverter a matriz Φ e multiplicá-la pela matriz Z :

$$\begin{aligned}\Phi\Omega_r &= Z \\ \Omega_r &= \Phi^{-1}Z\end{aligned}\quad (2.2.13)$$

2.2.3 Anisotropia em Redes RBF

Em alguns casos, observa-se nos dados amostrais uma tendência de variação maior em uma determinada direção. A Figura 10, mostrada na Seção 2.1.6.5, mostra um exemplo deste tipo de dado, onde horizontalmente os dados variam muito pouco quando comparados com a variação vertical. Nestes casos, somente a distância não é suficiente para descrever o grau de dependência espacial e a direção passa a ser importante também.

As funções de ativação, explicadas na Seção 2.2, utilizam a distância como parâmetro. A fórmula mais simples de distância é a Distância Euclidiana, mostrada na Equação (2.1.86). Utilizando essa distância na interpolação das Redes RBF, as funções de ativação dos neurônios possuirão ativações idênticas em qualquer direção, e essa característica é passada ao resultado final da interpolação. Funções de ativação desse tipo são chamadas de isotrópica. Funções de ativação que possuem variação conforme a direção são chamadas de anisotrópicas. A Figura 18 mostra uma função de ativação gaussiana isotrópica comparada com uma anisotrópica.

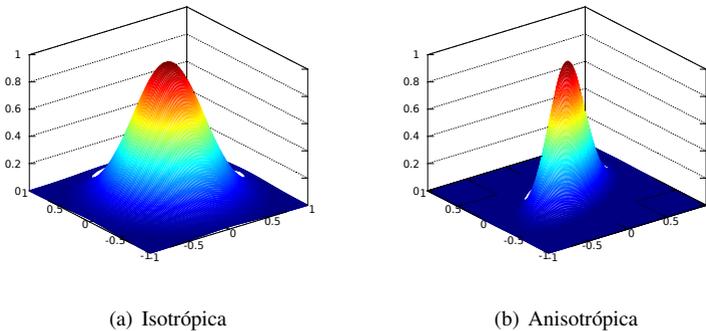


Figura 18: Exemplo de funções de ativação com duas dimensões.

A Figura 19 mostra os resultados de interpolações usando funções de ativação isotrópicas, que não mudam conforme a direção da distância, e anisotrópicas, que mudam. Também é mostrada uma imagem com todos os dados reais e os pontos usados como dados amostrais. Os marcadores rosa representam os centros da função de ativação de cada neurônio. Como os dados reais têm característica anisotrópica, a versão anisotrópica da função de ativação gerou uma interpolação melhor.

Para utilizar a versão anisotrópica, as Redes RBF utilizam uma matriz de covariância, que é uma matriz onde a diagonal principal contém a variância

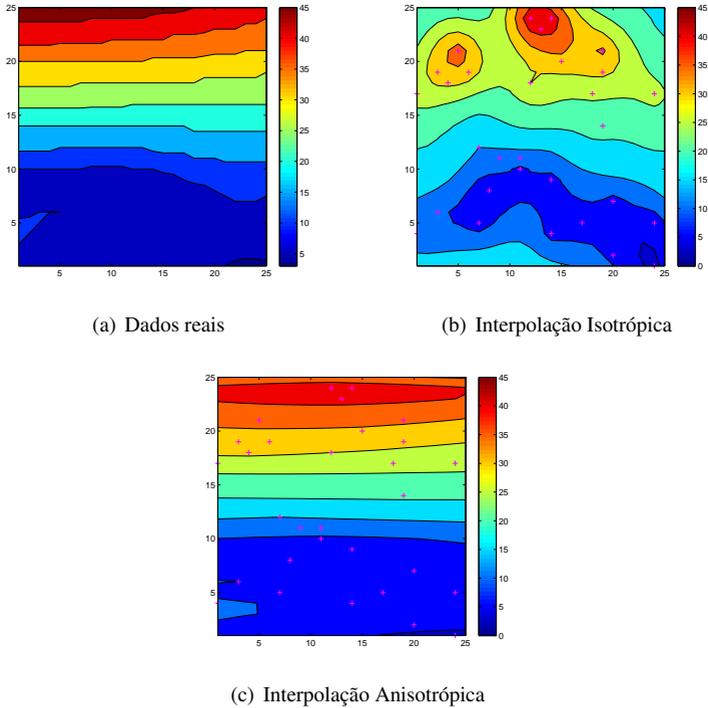


Figura 19: Exemplo de interpolação RBF com dados anisotrópicos.

de cada coordenada e as demais posições contém a covariância entre cada coordenada. A Equação (2.1.86) é aqui reescrita utilizando uma matriz de covariância que resulta na mesma distância Euclidiana:

$$\begin{aligned} d_x &= x_a - x_b \\ d_y &= y_a - y_b \end{aligned} \quad (2.2.14)$$

$$\text{distância}(a, b) = \sqrt{[d_x \quad d_y] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix}},$$

onde a e b são dois pontos diferentes, x_a e y_a são respectivamente as coordenadas horizontal e vertical do ponto a , assim como x_b e y_b são respectivamente as coordenadas horizontal e vertical do ponto b .

Para adicionar o efeito de anisotropia, basta modificar a matriz da Equação (2.2.14), que é chamada de matriz de covariância. A nova equação

é:

$$\begin{aligned}
 d_x &= x_a - x_b \\
 d_y &= y_a - y_b \\
 \text{distância}(a,b) &= \sqrt{\begin{bmatrix} d_x & d_y \end{bmatrix} \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix}} \quad (2.2.15) \\
 \text{sendo } c_x &> 0 \\
 c_y &> 0 \\
 -\sqrt{c_x c_y} &< c_v < \sqrt{c_x c_y},
 \end{aligned}$$

onde os novos parâmetros c_x , c_y e c_v precisam ser estimados. As restrições de valores possíveis para estes parâmetros são para garantir que o resultado da equação seja sempre um número real positivo. Esses parâmetros podem ser entendidos graficamente pela Figura 20. Quando parâmetro c_v vai se afastando do zero, maior a diferença entre as duas retas representadas pelo parâmetro. Porém, essa relação não é linear. O ângulo de anisotropia também é controlado pela variável c_v , o que torna difícil deduzir graficamente o seu valor. Os parâmetros c_x e c_y são mais fáceis de compreender, sendo apenas o quadrado da medida indicada no gráfico. Quando $c_x = c_y$ e $c_v = 0$, a função de ativação se torna isotrópica.

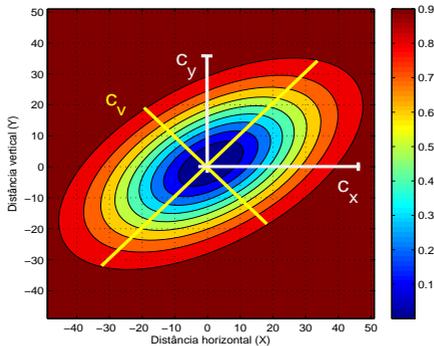


Figura 20: Parâmetros da função de ativação anisotrópica.

2.2.4 Algoritmo das Redes RBF Exatas

Os pseudocódigos das Redes RBF exatas estão descritos nos algoritmos 4, 5 e 6.

Dados:

Dados amostrais (n dados), representados pela matriz Z . O símbolo $Z[1]$ representa o valor do dado amostral no ponto 1.

Função de ativação ϕ já parametrizado.

Funções de Distância, MultiplicaMatrizes, MatrizInversa e MatrizTransposta.

Resultado:

Dados interpolados representados por R . O símbolo $R[j]$ representa o valor interpolado no ponto j .

início

para $a \leftarrow 1$ até n **faça**

para $b \leftarrow 1$ até n **faça**

$\Phi[a,b] \leftarrow \phi(\text{Distância}(a,b));$

fim

fim

$\Omega_r \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Phi), Z);$

para cada ponto j dos pontos a serem interpolados **faça**

para $a \leftarrow 1$ até n **faça**

$\Upsilon[a] \leftarrow \phi(\text{Distância}(a,j));$

fim

$R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_r), \Upsilon);$

fim

fim

Algoritmo 4: Redes RBF Exatas sem *bias*

Dados:

Dados amostrais (n dados), representados pela matriz Z . O símbolo $Z[1]$ representa o valor do dado amostral no ponto 1.

Função de ativação ϕ já parametrizado.

Funções de Distância, MultiplicaMatrizes, MatrizInversa e MatrizTransposta.

Resultado:

Dados interpolados representados por R . O símbolo $R[j]$ representa o valor interpolado no ponto j .

início

para $a \leftarrow 1$ **até** n **faça**

para $b \leftarrow 1$ **até** n **faça**

$\Phi[a, b] \leftarrow \phi(\text{Distância}(a, b));$

fim

$\Phi[a, n+1] \leftarrow 1;$

$\Phi[n+1, a] \leftarrow 1;$

fim

$\Phi[n+1, n+1] \leftarrow 0;$

$Z[n+1] \leftarrow 0;$

$\Omega_r \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Phi), Z);$

para cada ponto j **dos pontos** a **serem interpolados** **faça**

para $a \leftarrow 1$ **até** n **faça**

$\Upsilon[a] \leftarrow \phi(\text{Distância}(a, j));$

fim

$\Upsilon[n+1] \leftarrow 1;$

$R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_r), \Upsilon);$

fim

fim

Algoritmo 5: Redes RBF Exatas com *bias* constante

Dados:

Idem ao das Redes RBF Exatas sem *bias*, com a adição das coordenadas dos pontos, representados pelas letras X e Y . O símbolo $X[a]$ representa a coordenada horizontal do ponto a .

Resultado:

Idem ao das Redes RBF Exatas sem *bias*.

início

```

para  $a \leftarrow 1$  até  $n$  faça
  para  $b \leftarrow 1$  até  $n$  faça
     $\Phi[a, b] \leftarrow \phi(\text{Distância}(a, b));$ 
  fim
   $\Phi[a, n+1] \leftarrow 1;$ 
   $\Phi[a, n+2] \leftarrow X[a];$ 
   $\Phi[a, n+3] \leftarrow Y[a];$ 
   $\Phi[n+1, a] \leftarrow 1;$ 
   $\Phi[n+2, a] \leftarrow X[a];$ 
   $\Phi[n+3, a] \leftarrow Y[a];$ 
fim
para  $a \leftarrow n+1$  até  $n+3$  faça
  para  $b \leftarrow n+1$  até  $n+3$  faça
     $\Phi[a, b] \leftarrow 0;$ 
  fim
fim
para  $a \leftarrow n+1$  até  $n+3$  faça
   $Z[a] \leftarrow 0;$ 
fim
 $\Omega_r \leftarrow \text{MultiplicaMatrizes}(\text{MatrizInversa}(\Phi), Z);$ 
para cada ponto  $j$  dos pontos  $a$  serem interpolados faça
  para  $a \leftarrow 1$  até  $n$  faça
     $Y[a] \leftarrow \phi(\text{Distância}(a, j));$ 
  fim
   $Y[n+1] \leftarrow 1;$ 
   $Y[n+2] \leftarrow X[j];$ 
   $Y[n+3] \leftarrow Y[j];$ 
   $R[j] \leftarrow \text{MultiplicaMatrizes}(\text{MatrizTransposta}(\Omega_r), Y);$ 
fim
fim

```

Algoritmo 6: Redes RBF Exatas com *bias* variável

2.3 KRIGAGEM NEURAL E COMPARAÇÕES

A Krigagem e as Redes Neurais Artificiais já foram comparadas em diversos trabalhos. Nesta dissertação foram levantados alguns destes trabalhos a fim de verificar os resultados obtidos.

Um trabalho (SAKATA; ASHIDA; ZAKO, 2010) compara a Krigagem com Redes Neurais Hierárquicas, utilizando a função seno como dados de teste. Como a função é conhecida, foi possível calcular o erro de interpolação em todos os pontos. Conclui-se, no experimento realizado, que para poucos dados amostrais as Redes Neurais Hierárquicas têm menor erro, mas quando a quantidade de dados usados para interpolar vai aumentando, o erro de interpolação da Krigagem vai ficando cada vez menor, ao contrário das Redes Neurais Hierárquicas.

Outro trabalho (MATÍAS et al., 2004) compara os métodos de interpolação: Krigagem, Redes RBF, Redes MLP e Redes de Regularização. Após experimentos, este trabalho conclui que as Redes MLP geram resultados piores, em relação à suavidade e generalização da interpolação, que os demais métodos.

Em outro trabalho (BADEL; ANGORANI; PANAH, 2011) é comparada as Redes MLP com a Krigagem em um experimento relacionado a volume de depósito de ferro. Alguns dados amostrais não foram utilizados na interpolação, com o intuito de servirem para medir a correlação e o erro da interpolação. No experimento, a Krigagem gerou maior correlação com os dados separados, e menor erro, mas acabou repetindo o erro contido nos dados amostrais que foram utilizados para interpolar. Apesar do maior erro, as Redes MLP reproduziram melhor as características desejadas.

Um artigo (COSTA; PRONZATO; THIERRY, 1999) compara a Krigagem com as Redes RBF e conclui que ambos compartilham a mesma estrutura. Após experimentos utilizando a função matemática sinc , definida por $\text{seno}(x)/x$, conclui-se que a Krigagem é mais precisa que as Redes RBF para interpolar esta função, já que foi possível verificar o erro de cada interpolador por ser uma função conhecida.

Em outro artigo (SATHYANARAYANAMURTHY; CHINNAM, 2009) é feita uma comparação entre os métodos de interpolação: Krigagem, Redes RBF e Máquina de Vetores de Suporte, mais conhecido por SVM ou *Support Vector Machine*. Concluiu-se, após experimentos, que geralmente a Krigagem obtém melhores resultados, mas os outros dois métodos devem ser considerados quando se pretende evitar sobreajuste, também chamado de *overfitting*, que é o efeito indesejado de interpolar muito bem até o erro dos dados amostrais, gerando uma interpolação pior.

Em um trabalho (YERRABOLU et al., 2012) é criada uma variação

das Redes MLP e este modelo é comparado com Redes RBF, IDW, interpolação com polinômios, Krigagem e Redes MLP, com o objetivo de encontrar a distribuição geográfica de um agente cancerígeno. Após experimentos, o trabalho conclui que o modelo proposto obteve melhores resultados.

Em outro trabalho (LI et al., 2013) são comparados os seguintes métodos de interpolação: Redes RBF, Krigagem e Regressão Linear. Após experimento sobre a distribuição de matéria orgânica do solo, concluiu-se que as Redes RBF obtiveram melhores resultados.

Um artigo (AKCIN; CELIK, 2013) interpola um geóide local, para aplicação de cálculo de altitude, utilizando os métodos Redes MLP e Krigagem. Após experimentos, conclui-se que as Redes MLP obtiveram melhores resultados.

Em outro artigo (CERESETTI et al., 2012) é comparada a interpolação das Redes MLP e da Krigagem para chuvas extremas. Conclui-se, após experimentos, que os métodos possuem resultados comparáveis, mas a Krigagem obtém melhores resultados quando os dados são bem correlacionados espacialmente.

Outro artigo (ZHANG; GOVINDARAJU, 2000) comparou as Redes RBF e a Krigagem. Concluiu-se que as Redes RBF interpolam mais suavemente segundo experimentos.

Existem trabalhos que criam novos métodos mesclando as Redes Neurais Artificiais com a Krigagem, o que também é chamado de Krigagem Neural. É importante verificar os resultados obtidos, sobre mesclar os métodos, em outros trabalhos, a fim de guiar este estudo. Os trabalhos descritos a partir daqui são relacionados à Krigagem Neural.

Um trabalho de (LIN; CHEN, 2004) cria um novo método inserindo o semivariograma da Krigagem nas funções de ativação das Redes RBF. O resultado, segundo um experimento realizado, mostra que o novo método é melhor que as Redes RBF e a Krigagem. O motivo provável é que nas Redes RBF a quantidade de neurônios é arbitrária e a posição do centro das RBF também, diferente da Krigagem, onde todos os cálculos são feitos a partir da distância dos dados amostrais.

Outro trabalho (ZHENG et al., 2009) compara os métodos: Krigagem, as Redes MLP e um método de Krigagem Neural, que primeiro interpola usando Redes MLP e depois ajusta o erro com a Krigagem. Segundo experimentos, a Krigagem Neural obtém melhores resultados que a Krigagem e as Redes MLP.

Em outro trabalho (DEMYANOV et al., 1998) é criado outro método de Krigagem Neural que utiliza Redes MLP e, em seguida, corrige o erro de interpolação com Krigagem. O trabalho conclui que a Krigagem Neural é mais fácil de utilizar que a Krigagem porque a tendência pode ser calculada

pelas Redes MLP ao invés do usuário procurar um modelo adequado para a tendência dos dados.

Um artigo (DEMYANOV et al., 2001) cria outro método de Krigagem Neural que realiza a interpolação utilizando um método chamado Interpolação de *Wavelet*, descrito no trabalho referenciado, para depois corrigir o erro através da Krigagem. Além disso, é feita uma comparação com outro método de Krigagem Neural, que utiliza Redes MLP ao invés de Interpolação de *Wavelet*. Conclui-se, através de experimentos, que a interpolação de *Wavelet* é superior quando os dados são anisotrópicos.

Em outro artigo (CELLURA et al., 2008) é proposta uma Krigagem Neural onde, primeiramente, é feita uma interpolação utilizando Redes MLP para depois corrigir o erro utilizando Krigagem. Conclui-se, pelo experimento, que os resultados não foram satisfatórios devido a pouca quantidade de dados amostrais, e que o método criado é mais simples e fácil de implementar do que o método utilizado tradicionalmente no problema abordado pelo artigo de interpolação de ventos.

Finalmente, outro trabalho (KOIKE; MATSUDA; GU, 2001) cria um novo tipo de Krigagem Neural, modificando o treinamento das Redes MLP para que considere a diferença do semivariograma teórico da Krigagem em relação ao semivariograma obtido com as Redes MLP. Após experimentos realizados, conclui-se que a Krigagem Neural proposta obtém melhores resultados que as Redes MLP e a Krigagem.

3 COMPARAÇÃO ANALÍTICA

Apesar de haver uma boa quantidade de trabalhos que fazem uma comparação entre os métodos de Krigagem e Redes Neurais Artificiais para aplicações de interpolação espacial, poucos trabalhos fazem uma comparação analítica visando estudar os métodos sob um mesmo arcabouço formal. Este tipo de comparação visa identificar as diferenças para criar um novo método que incorpora as vantagens dos dois métodos.

Assim, neste capítulo, primeiramente, é feito uma comparação dos métodos de interpolação, como o objetivo de compará-los analiticamente, utilizando o mesmo modelo para o semivariograma e a função de ativação.

Em seguida, é feita uma comparação analítica dos parâmetros utilizados na função de ativação das Redes RBF e os parâmetros utilizados no semivariograma, a fim de encontrar uma relação algébrica entre eles que gerem o mesmo resultado. Desta forma, é demonstrada a equivalência dos parâmetros.

Finalmente, é calculada a complexidade dos métodos analisados, a fim de verificar qual método é mais eficiente computacionalmente.

3.1 COMPARAÇÃO DOS MÉTODOS DE INTERPOLAÇÃO

Os métodos de interpolação são comparados aqui com o objetivo de evidenciar as reais diferenças entre eles. Para isso, sempre que possível, é dado o mesmo tratamento e parâmetros para ambos os métodos.

O modelo da função de ativação das Redes RBF utilizada será igual ao modelo da função de semivariograma. Apesar de serem mais usadas funções decrescentes para a função de ativação das Redes RBF, elas podem ser crescentes também, conforme mostrado na Seção 2.2. Da mesma forma, o semivariograma é uma função crescente, mas pode ser uma função decrescente, sendo chamado de covariograma neste caso, que é equivalente ao semivariograma com limitações, como explicado na Seção 2.1.7. A comparação será feita considerando o mesmo modelo e parâmetros para o semivariograma e a função de ativação. Apesar do processos de estimar os parâmetros do semivariograma e Redes RBF serem diferentes, serão considerados os mesmos parâmetros a fim de facilitar a comparação. Na Seção 3.2 é feita uma comparação analítica entre os parâmetros utilizados na função de ativação das Redes RBF e do semivariograma. Dessa forma, será assumido que:

$$\begin{array}{c} \text{Krigagem - semivariograma} \\ \underbrace{\quad\quad\quad}_{\gamma} \end{array} = \begin{array}{c} \text{Redes RBF - função de ativação} \\ \underbrace{\quad\quad\quad}_{\phi} \end{array}, \quad (3.1.1)$$

onde γ é o modelo de semivariograma e ϕ é a função de ativação. Essa assunção será considerada na conclusão da comparação.

É utilizada na comparação a Krigagem Universal, já que a Krigagem Ordinária e Simples são casos específicos da Universal, como mostrado na Seção 2.1.3.

Para as Redes RBF, é usada a variação mais próxima à Krigagem Universal: a que utiliza um polinômio com as coordenadas verticais e horizontais como *bias*, descrita na Seção 2.2.2.

A Equação (2.1.4), que define a interpolação da Krigagem, pode ser reescrita em forma de multiplicação matricial:

$$\begin{aligned}
 z_j &= \sum_i \omega_i z_i \\
 &= \overbrace{[\omega_1 \quad \omega_2 \quad \dots \quad \omega_n \quad \lambda_c \quad \lambda_x \quad \lambda_y]}^{\mathbf{\Omega}_k^T} \underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \dots \\ z_n \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{Z}} \\
 &= \mathbf{\Omega}_k^T \mathbf{Z}, \tag{3.1.2}
 \end{aligned}$$

onde $\mathbf{\Omega}_k$ é a matriz de pesos e parâmetros a serem estimados e \mathbf{Z} é a matriz com os dados amostrais. O motivo para ter sido colocado zeros para anular os parâmetros λ é simplificar a expressão em função de variáveis que aparecem em outras equações de forma a facilitar a comparação. O importante é que esta equação é equivalente à original.

De forma semelhante, a Equação (2.2.11), que define a interpolação

das Redes RBF, pode ser reescrita como:

$$z_j = \lambda_c + x_j \lambda_x + y_j \lambda_y + \sum_i \phi_{ij} \omega_i$$

$$= \underbrace{[\omega_1 \quad \omega_2 \quad \dots \quad \omega_n \quad \lambda_c \quad \lambda_x \quad \lambda_y]}_{\mathbf{\Omega}_r^T} \underbrace{\begin{bmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \dots \\ \gamma_{nj} \\ 1 \\ x_j \\ y_j \end{bmatrix}}_{\mathbf{r}}$$

$$= \mathbf{\Omega}_r^T \mathbf{r}, \quad (3.1.3)$$

onde $\mathbf{\Omega}_r$ é a matriz contendo os pesos e parâmetros a serem estimados e \mathbf{r} é a matriz com a ativação de cada neurônio para o ponto a ser interpolado. A função ϕ foi substituída por γ devido à suposição de equivalência, explicada na Equação (3.1.1).

As equações (2.1.64) e (2.2.13), que são usadas para calcular os pesos e parâmetros na Krigagem e nas Redes RBF, estão reescritas abaixo para facilitar a análise:

$$\text{Krigagem: } \mathbf{\Omega}_k = \mathbf{\Gamma}^{-1} \mathbf{Y} \quad (3.1.4)$$

$$\text{Redes RBF: } \mathbf{\Omega}_r = \mathbf{\Gamma}^{-1} \mathbf{Z}, \quad (3.1.5)$$

onde a matriz $\mathbf{\Phi}$ foi substituída por $\mathbf{\Gamma}$ devido à suposição de equivalência entre os modelos ϕ e γ , explicada na Equação (3.1.1).

A matriz $\mathbf{\Gamma}$ é simétrica porque o resultado da função de ativação ou semivariograma entre o ponto a e b é o mesmo que entre b e a . Uma matriz simétrica é igual a sua transposta:

$$\mathbf{\Gamma} = \mathbf{\Gamma}^T \quad (3.1.6)$$

Utilizando estas informações, as equações usadas para interpolar serão reescritas para facilitar a comparação e verificar a diferença entre os métodos. As operações executadas estão explicadas na coluna da direita.

- **Krigagem:**

Equação	Explicação
$z_j = \mathbf{\Omega}_k^T \mathbf{Z}$	da Eq. (3.1.2)
$= (\mathbf{\Gamma}^{-1} \mathbf{Y})^T \mathbf{Z}$	da Eq. (3.1.4)
$= \mathbf{Y}^T (\mathbf{\Gamma}^{-1})^T \mathbf{Z}$	$(AB)^T = B^T A^T$
$= \mathbf{Y}^T (\mathbf{\Gamma}^T)^{-1} \mathbf{Z}$	$(A^{-1})^T = (A^T)^{-1}$
$= \mathbf{Y}^T \mathbf{\Gamma}^{-1} \mathbf{Z}$	da Eq. (3.1.6)

- **Redes RBF:**

Equação	Explicação
$z_j = \mathbf{\Omega}_r^T \mathbf{Y}$	da Eq. (3.1.3)
$= \mathbf{Y}^T \mathbf{\Omega}_r$	$A^T B = B^T A$
$= \mathbf{Y}^T \mathbf{\Gamma}^{-1} \mathbf{Z}$	da Eq. (3.1.5)

Portanto:

$$\underbrace{\mathbf{Y}^T \mathbf{\Gamma}^{-1} \mathbf{Z}}_{\text{Krigagem}} = \underbrace{\mathbf{Y}^T \mathbf{\Gamma}^{-1} \mathbf{Z}}_{\text{Redes RBF}}, \quad (3.1.7)$$

o que mostra que as Redes RBF e a Krigagem são equivalentes em termos de modelagem, mas o processo de estimação dos parâmetros do semivariograma e função de ativação diverge.

As Redes RBF e a Krigagem possuem a mesma estrutura. Ambas podem ser descritas em termos de \mathbf{Y} , $\mathbf{\Gamma}$ e \mathbf{Z} , como mostrado na Equação (3.1.7). Por isso, elas podem compartilhar as melhorias feitas em cada método. Por exemplo, as Redes RBF não possuem um método para estimar o intervalo de predição que a Krigagem possui. Se a função de ativação for ajustada pelo semivariograma experimental, esse intervalo pode ser calculado. Outro exemplo é que normalmente as Redes RBF não utilizam um polinômio na sua interpolação, entretanto podem ser usados os mesmos polinômios da Krigagem. Quanto a Krigagem, pode-se utilizar a técnica das Redes RBF de agrupar os dados amostrais espacialmente próximos, tornando a interpolação mais suave.

Porém, a utilização de um procedimento de um método em outro deve ser feita com cautela. Alguns destes procedimentos se baseiam nas premissas de cada método e podem não fazer sentido utilizá-la no outro método sem esse cuidado. Por exemplo, para estimar o erro de interpolação nas Redes RBF é necessário que a função de ativação seja ajustada pelo semivariograma experimental, e que seja um semivariograma válido, de acordo com o exposto na Seção 2.1.

Uma variação da Krigagem, proposta em alguns trabalhos (SRINIVASAN; DURAI SWAMI; MURTUGUDDE, 2011; TROCHU, 1993), faz modificações no método que a tornam mais rápida, utilizando a mesma forma de

interpolares das Redes RBF.

3.2 COMPARAÇÃO ENTRE OS PARÂMETROS

Nesta seção são comparados os parâmetros do semivariograma com os parâmetros da função de ativação das Redes RBF em termos de efeito dos parâmetros no resultado do modelo. A importância desta comparação é verificar se é sempre possível converter uma função de ativação, com os mesmos resultados, em um semivariograma.

Os parâmetros do semivariograma, no caso mais completo de distância anisotrópica, são: *alcance*, *pepita*, *limiar*, fator de anisotropia r e ângulo α , conforme descritos na Seção 2.1.6.5. Os parâmetros da equação da função de ativação, também com distância anisotrópica, são: c_x , c_y e c_v , que são mostrados na Seção 2.2.3. Os outros parâmetros da função de ativação podem ser fixados para que apenas os parâmetros descritos precisem ser ajustados, de forma que c_x e c_y representem o quadrado da distância horizontal e vertical, como ilustrado na Figura 20.

Aqui é feita a transformação dos parâmetros do semivariograma em parâmetros da função de ativação das Redes RBF e vice-versa. Dessa forma, prova-se a equivalência entre os parâmetros.

O parâmetro *pepita* é inexistente na função de ativação, e ele será considerado zero para que seu efeito seja anulado. O parâmetro *limiar* também não existe na função de ativação, e ele é considerado 1, mas qualquer valor diferente de zero poderia ser considerado. Os parâmetros *pepita* e *limiar* não alteram o resultado da interpolação da Krigagem, conforme demonstrado na Seção 2.1.6.2. Eles servem apenas como base para que o modelo de semivariograma seja bem ajustado ao semivariograma experimental, porém sem efeito real na interpolação. Portanto, mesmo que não existam na função de ativação, eles não acrescentam informação útil ao modelo.

Resta apenas a conversão dos parâmetros *alcance*, r e α em c_x , c_y e c_v e vice versa.

Para facilitar, as equações de distância anisotrópicas de cada método serão reescritas aqui. Para o semivariograma é usado:

$$\begin{aligned} h_x &= \frac{1}{r} ((x_a - x_b) \cos \alpha - (y_a - y_b) \sin \alpha) \\ h_y &= (y_a - y_b) \cos \alpha + (x_a - x_b) \sin \alpha \quad (2.1.88) \\ \text{distância}(a, b) &= \sqrt{h_x^2 + h_y^2} \end{aligned}$$

Para a função de ativação é usado:

$$\begin{aligned}d_x &= x_a - x_b \\d_y &= y_a - y_b \\ \text{distância}(a, b) &= \sqrt{[d_x \quad d_y] \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} d_x \\ d_y \end{bmatrix}} \quad (2.2.15)\end{aligned}$$

$$\text{sendo } c_x > 0$$

$$c_y > 0$$

$$-\sqrt{c_x c_y} < c_v < \sqrt{c_x c_y}$$

Primeiro será feita a conversão dos parâmetros *alcance*, r e α em c_x , c_y e c_v . Os valores de c_x e c_y representam o alcance, ou seja, a distância onde o parâmetro se torna o valor máximo, na coordenada x e y , respectivamente. Portanto, através das funções *seno* e *coseno*, e através dos parâmetros *alcance*, r e α , é possível calcular seus valores, considerando o significado dos parâmetros:

$$c_x = (\text{alcance} \cdot r \cdot \sin \alpha)^2 + (\text{alcance} \cdot \cos \alpha)^2 \quad (3.2.1)$$

$$c_y = (\text{alcance} \cdot r \cdot \cos \alpha)^2 + (\text{alcance} \cdot \sin \alpha)^2 \quad (3.2.2)$$

O valor de c_v exige mais operações. A ideia é escrever duas equações equivalentes que utilizam ângulos diferentes de forma que o c_v possa ser isolado. A Equação (2.2.15) utiliza as distâncias d_x e d_y entre dois pontos. Definiremos d_x e d_y utilizando uma distância m e um ângulo α , resultando em:

$$d_{x1} = m \cdot \sin \alpha \quad (3.2.3)$$

$$d_{y1} = m \cdot \cos \alpha \quad (3.2.4)$$

As distâncias d_{x1} e d_{y1} estão na direção em que a elipse da função de ativação é maior. Agora é considerado um novo ponto, cujo valor da função de ativação será o mesmo, mas com ângulo diferente. O novo ângulo será onde a elipse é menor, a 90 graus da maior variação. Como a distância diminuirá proporcional ao fator de anisotropia r , a nova distância, chamada de d_{x2} e

d_{y2} , é:

$$d_{x2} = \frac{m \cdot \sin(\alpha + \frac{\pi}{2})}{r} \quad (3.2.5)$$

$$d_{y2} = \frac{m \cdot \cos(\alpha + \frac{\pi}{2})}{r} \quad (3.2.6)$$

Essas duas distâncias resultam na mesma ativação. Portanto, pela Equação (2.2.15):

$$\begin{bmatrix} d_{x1} & d_{y1} \end{bmatrix} \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} d_{x1} \\ d_{y1} \end{bmatrix} = \begin{bmatrix} d_{x2} & d_{y2} \end{bmatrix} \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} d_{x2} \\ d_{y2} \end{bmatrix} \quad (3.2.7)$$

Agora, substitui-se o d_{x1} , d_{y1} , d_{x2} e d_{y2} :

$$\begin{bmatrix} m \cdot \sin \alpha & m \cdot \cos \alpha \end{bmatrix} \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} m \cdot \sin \alpha \\ m \cdot \cos \alpha \end{bmatrix} = \\ \begin{bmatrix} \frac{m \cdot \sin(\alpha + \frac{\pi}{2})}{r} & \frac{m \cdot \cos(\alpha + \frac{\pi}{2})}{r} \end{bmatrix} \begin{bmatrix} c_x & c_v \\ c_v & c_y \end{bmatrix} \begin{bmatrix} \frac{m \cdot \sin(\alpha + \frac{\pi}{2})}{r} \\ \frac{m \cdot \cos(\alpha + \frac{\pi}{2})}{r} \end{bmatrix} \quad (3.2.8)$$

Pode-se simplificar essa expressão até isolar o c_v :

Se α é múltiplo de $\frac{\pi}{2}$:

$$c_v = 0$$

senão:

$$c_v = \frac{\csc(2\alpha)}{r^2 + 1} ((c_x r^2 - c_y) \sin^2 \alpha + (c_y r^2 - c_x) \cos^2 \alpha) \quad (3.2.9)$$

Essa equação não está definida para α múltiplo de $\frac{\pi}{2}$, já que a função cossecante tende ao infinito nesses pontos. Neste caso, $c_v = 0$, pois ângulos múltiplo de $\frac{\pi}{2}$ são verticais e horizontais, onde somente os parâmetros c_x e c_y já definem a anisotropia.

Agora será feita a conversão dos parâmetros c_x , c_y e c_v em *alcance*, r e α , que é um pouco mais complexa.

Primeiro, o caso mais simples, onde $c_v = 0$, o ângulo α pode ser horizontal ou vertical, ou seja, 0 ou $\frac{\pi}{2}$. Além disso, se $c_x = c_y$ e $c_v = 0$ então a distância é isotrópica e não importa o valor do parâmetro α . Os parâmetros

podem ser escritos como:

se $c_v = 0$ então:

$$r = \sqrt{\frac{\max(c_y, c_x)}{\min(c_y, c_x)}} \quad (3.2.10)$$

$$\text{alcance} = \sqrt{\min(c_y, c_x)} \quad (3.2.11)$$

$$\alpha = \begin{cases} \frac{\pi}{2} & \text{se } c_y > c_x \\ 0 & \text{se } c_y < c_x \\ \text{qualquer} & \text{se } c_y = c_x \end{cases} \quad (3.2.12)$$

No caso mais complexo, quando $c_v \neq 0$, o r é isolado na relação $\frac{c_x}{c_y}$, obtidos pelas equações (3.2.1) e (3.2.2):

$$\frac{c_x}{c_y} = \frac{(\text{alcance} \cdot r \cdot \sin \alpha)^2 + (\text{alcance} \cdot \cos \alpha)^2}{(\text{alcance} \cdot r \cdot \cos \alpha)^2 + (\text{alcance} \cdot \sin \alpha)^2} \quad (3.2.13)$$

Após isolar o r , o *alcance* acaba sendo anulado, o que resulta na seguinte equação:

$$r = \sqrt{\frac{c_y \cdot \cos^2 \alpha - c_x \cdot \sin^2 \alpha}{c_x \cdot \cos^2 \alpha - c_y \cdot \sin^2 \alpha}} \quad (3.2.14)$$

Essa equação ainda possui duas variáveis não conhecidas: r e α . Agora, o r é isolado na Equação (3.2.9), resultando em

$$r = \sqrt{\frac{-c_x \cdot \cot \alpha - c_y \cdot \tan \alpha - 2c_v}{-c_x \cdot \tan \alpha - c_y \cdot \cot \alpha + 2c_v}} \quad (3.2.15)$$

Igualando as equações (3.2.14) e (3.2.15), sobra apenas a variável α , que pode ser isolada:

$$\alpha = \frac{-c_v}{|c_v|} \left(\pi - \arccos \left(\sqrt{\frac{(c_x - c_y) \sqrt{m + m}}{2m}} \right) \right) \quad (3.2.16)$$

sendo $m = 4c_v^2 + (c_x - c_y)^2$

Agora que o ângulo está definido somente pelos parâmetros do semi-variograma, é possível calcular o r utilizando a Equação (3.2.15), já que se possui as variáveis necessárias c_x , c_y , c_v e α .

Finalmente, para calcular o *alcance* basta isolá-lo da Equação (3.2.1):

$$\text{alcance} = \sqrt{\frac{c_y}{r^2 \sin^2 \alpha + \cos^2 \alpha}}, \quad (3.2.17)$$

onde as variáveis r e α foram calculadas pelas equações (3.2.15) e (3.2.16).

Foram feitos muitos testes transformando os parâmetros de um método para o outro para verificar que os cálculos estão corretos.

Como foi possível definir os parâmetros do semivariograma a partir dos parâmetros da função de ativação e vice-versa, está demonstrado que ambos os modelos são equivalentes. A função de ativação não tem os parâmetros *pepita* e *limiar*, mas eles não afetam a interpolação, como demonstrado na Seção 2.1.6.2.

3.3 COMPARAÇÃO DE COMPLEXIDADE

Nesta seção é mostrado o estudo da complexidade computacional dos dois métodos usando a notação “Big O”, que descreve o limite do comportamento da função quando os argumentos tendem ao infinito (KNUTH, 1968). A complexidade é uma função que calcula a quantidade de operações envolvidas em um método com relação à quantidade de argumentos. Ela foi descrita em função dos parâmetros:

- Complexidade do semivariograma γ ou função de ativação ϕ : f
- Quantidade de amostras (Krigagem) ou neurônios (Redes RBF): n

A complexidade envolvida em cada ponto a ser interpolado na Krigagem é descrita partindo das equações (2.1.64) e (3.1.2):

$$\begin{aligned} \text{complexidade(Krigagem)} &= \underbrace{O(nf)}_{\mathbf{Y}} + \underbrace{O(n^2)}_{\mathbf{\Gamma}^{-1}\mathbf{Y}} + \underbrace{O(n)}_{\mathbf{\Omega}_k^T \mathbf{Z}} \\ &= O(n^2 + n(f + 1)). \end{aligned} \quad (3.3.1)$$

A primeira parte mostra que a matriz \mathbf{Y} precisa ser calculado usando o semivariograma. Como a Krigagem precisa inverter a matriz $\mathbf{\Gamma}$ somente uma vez e reutilizar o resultado para a interpolação em todos os pontos, a etapa $\mathbf{\Gamma}^{-1}\mathbf{Y}_j$ é apenas a multiplicação de uma matriz n por n e uma matriz 1 por n . Após os pesos calculados, basta realizar a operação $\mathbf{\Omega}_k^T \mathbf{Z}$, que é apenas uma multiplicação entre duas matrizes n por 1.

Para medir a complexidade das Redes RBF é usada a Equação (3.1.3):

$$\begin{aligned} \text{complexidade(RedesRBF)} &= \underbrace{O(nf)}_{\mathbf{r}} + \underbrace{O(n)}_{\mathbf{\Omega}_r^T \mathbf{r}} \\ &= O(n(f+1)) \end{aligned} \quad (3.3.2)$$

A primeira parte mostra que a matriz \mathbf{Y} precisa ser calculado usando a função de ativação. Como as Redes RBF precisam calcular os pesos $\mathbf{\Omega}_r$ somente uma vez e reutilizá-los em todos os pontos a serem interpolados, resta somente a etapa $\mathbf{\Omega}_r^T \mathbf{Y}$, que é apenas uma multiplicação entre duas matrizes n por 1.

É notável a diferença de desempenho entre a Krigagem e as Redes RBF, sendo a Krigagem mais complexa e, portanto, menos eficiente computacionalmente:

$$\underbrace{O(n^2 + n(f+1))}_{\text{Krigagem}} > \underbrace{O(n(f+1))}_{\text{Redes RBF}}. \quad (3.3.3)$$

4 EXPERIMENTOS

Alguns experimentos foram feitos para verificar as complexidades dos métodos, calculadas no Capítulo 3, bem como demonstrar de maneira prática a equivalência matemática dos dois métodos. Para isso, tanto o algoritmo de Krigagem quanto o algoritmo das Redes RBF foram implementados na linguagem Octave (OCTAVE, 2013) com a menor diferença possível entre as implementações para que os resultados sejam mais precisos e facilitar a comparação entre os métodos. Os códigos fontes podem ser obtidos entrando em contato com o autor.

No primeiro experimento é verificado o impacto do aumento da quantidade de dados amostrais no tempo de resposta das interpolações. Espera-se que esse aumento seja muito maior na Krigagem, já que a complexidade é maior.

No segundo experimento é verificado o impacto do aumento da quantidade de dados interpolados no tempo de resposta das interpolações. Agora, espera-se que o impacto relativo seja igual nos dois casos, já que mantendo a complexidade constante, por causa da mesma quantidade de dados amostrais, o aumento causará a multiplicação do número de pontos interpolados por esta constante, o que será proporcional nos dois casos.

No terceiro experimento são feitas interpolações utilizando Krigagem através de um software comercial chamado GS+, produzido pela empresa Gamma Design Software (GAMMADESIGN, 2013). O objetivo é constatar se o software utiliza a Krigagem na forma original ou se utiliza alguma forma otimizada, que melhore a complexidade calculada neste trabalho. Para isso, é verificado se o impacto de diferentes quantidades de dados amostrais e dados interpolados no tempo de processamento é conforme o esperado pelo cálculo de complexidade ou se é mais eficiente.

O quarto experimento tem o mesmo objetivo do terceiro, mas utiliza uma biblioteca chamada mGstat para MATLAB (MATHWORKS, 2013), desenvolvida por Thomas Mejer Hansen (HANSEN, 2013).

4.1 EXPERIMENTO DE AUMENTO DE DADOS AMOSTRAIS

Neste experimento é verificado o impacto do aumento da quantidade de dados amostrais no tempo de resposta das interpolações. Em um primeiro caso são utilizados 50 dados amostrais e, em outro, 467 dados amostrais, para estimar aproximadamente 10 mil novos pontos, que é uma grade 100 por 100, e formar um mapa de interpolação.

Foi utilizado modelo Gaussiano como função de ativação e semivariograma, com os parâmetros equivalentes para a Krigagem e Redes RBF. Foi utilizada a Krigagem Ordinária e a versão equivalente das Redes RBF, com *bias* constante, descrita na Seção 2.2.2.

Os dados foram retirados de um depósito de minério de ferro simulado (CLARK; HARPER, 2000) e de amostras de precipitação (DUBOIS, 1998). A Figura 21 mostra que os resultados são idênticos na interpolação do volume de minério de ferro e a Figura 22 mostra o mesmo para a interpolação que utiliza dados de precipitação. Esses resultados idênticos eram esperados pela demonstração teórica. Apesar de não ser possível visualizar com precisão os valores nas figuras, os resultados gerados pelos dois algoritmos são idênticos em todos os 10 mil valores interpolados. Os marcadores rosa demonstram a posição dos dados amostrais.

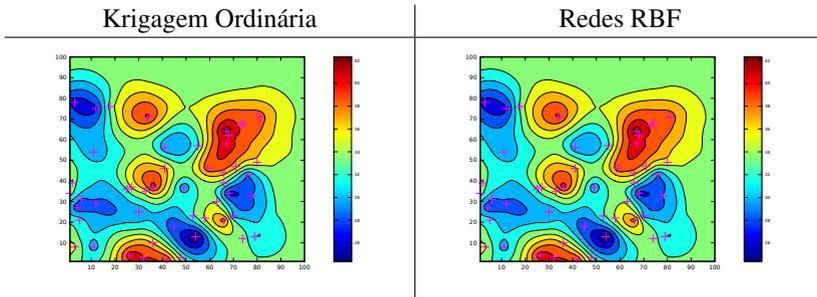


Figura 21: Experimento usando 50 dados amostrais de volume de minério de ferro.

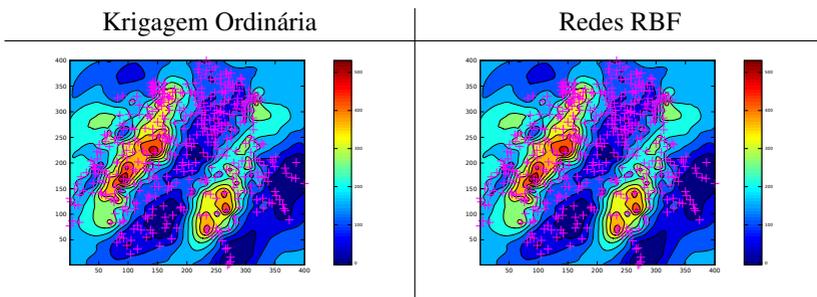


Figura 22: Experimento usando 467 dados amostrais de precipitação.

Os resultados em tempo de processamento, para uma comparação relativa entre os dois métodos, estão na Tabela 1.

Tabela 1: Tempo de processamento para interpolação de 10 mil pontos

Método	50 dados	467 dados	Aumento no tempo
Krigagem	2,8 segundos	1415 segundos	500 vezes
Redes RBF	1,2 segundos	2 segundos	2 vezes

Há uma diferença significativa entre os dois métodos quando a quantidade de dados amostrais aumenta. Essa diferença era esperada pelo cálculo de complexidade feito na Seção 3.3, o que comprova o que já era esperado pela teoria.

4.2 EXPERIMENTO COM AUMENTO DE PONTOS INTERPOLADOS

Neste experimento é verificado o impacto do aumento da quantidade de dados interpolados sem alterar a quantidade de dados amostrais no tempo de resposta das interpolações. São utilizados 50 dados amostrais de volume de minério de ferro (CLARK; HARPER, 2000) para interpolar, em um primeiro caso, 10 mil pontos, que é uma grade de 100 por 100, e em um segundo caso, 160 mil pontos, que é uma grade 400 por 400.

Foi utilizado modelo Gaussiano como função de ativação e semivariograma, com os parâmetros equivalentes. Foi utilizada a Krigagem Ordinária e sua versão equivalente das Redes RBF, descrita na Seção 2.2.2, como no experimento anterior.

Os resultados em tempo de processamento, para uma comparação relativa entre os dois métodos, são mostrados na Tabela 2.

Tabela 2: Tempo de processamento para interpolação de 160 mil pontos

Método	10 mil	160 mil	Aumento no tempo
Krigagem:	2,8 segundos	45 segundos	16 vezes
Redes RBF:	1,2 segundos	19 segundos	16 vezes

Apesar de a Krigagem ter tido um pior desempenho computacional, como esperado pela complexidade, o aumento de pontos interpolados causou um aumento relativamente igual nos dois casos: 16 vezes. O cálculo de complexidade não inclui a quantidade de pontos interpolados mas a quantidade de dados amostrais. Se as variáveis que compõem a complexidade forem constantes, era esperado que o aumento fosse relativamente igual nos dois casos, pois toda a complexidade será um valor constante multiplicado pela quantidade de dados interpolados nos dois casos.

4.3 EXPERIMENTOS COM SOFTWARE COMERCIAL

Neste experimento são feitas interpolações de Krigagem utilizando o software comercial GS+ com o objetivo de constatar se o software utiliza a Krigagem na forma original ou se utiliza alguma forma otimizada, que melhore a complexidade calculada. Para isso, é verificado se o impacto de diferentes quantidades de dados amostrais e dados interpolados no tempo de processamento é como previsto pela complexidade calculada ou mais rápido.

Este software limita o uso de no máximo 64 pontos amostrais ao redor do dado, provavelmente para evitar que o tempo de processamento seja muito grande. Por isso, foram feitas 4 interpolações utilizando os dados de um depósito de minério de ferro simulado (CLARK; HARPER, 2000). Em duas interpolações foram usados 50 dados amostrais e, nas outras duas, 25 dados amostrais. Além disso, em duas interpolações foram estimados aproximadamente 10 mil pontos, que é uma grade de 100 por 100, enquanto nas outras duas foram estimados aproximadamente 40 mil pontos, que é uma grade de 200 por 200.

Tabela 3: Tempo de processamento para a Krigagem em software comercial

Krigagem no GS+	10 mil pontos interp.	40 mil pontos interp.
25 dados amostrais	18 segundos	60 segundos
50 dados amostrais	31 segundos	80 segundos

Os resultados, mostrados na Tabela 3, não apresentaram o aumento obtido no segundo experimento, descrito na Seção 4.2. O motivo provável é que o software utilize otimizações que permitem a execução mais eficiente do método de Krigagem. Entretanto, por ser um sistema comercial, o fabricante não divulga o código e tão pouco esclarece quais otimizações são feitas no algoritmo da Krigagem. A limitação de utilizar no máximo 64 dados amostrais ao mesmo tempo para interpolar um ponto também faz com que não seja necessário resolver a inversa de uma matriz muito grande, o que melhora o desempenho.

4.4 EXPERIMENTOS COM BIBLIOTECA DE KRIGAGEM

Para verificar se outras implementações bastante populares de software de Geoestatística também fazem otimizações no algoritmo da Krigagem ou se usam o algoritmo padrão descrito no capítulo 2, fizemos um experimento com uma biblioteca de Geoestatística chamada mGstat, que se propõe

a ser a biblioteca padrão de Geoestatística para o MATLAB. Foi repetido o experimento de aumento de dados amostrais, descrito na Seção 4.1, mas utilizando somente a Krigagem através da biblioteca citada.

Tabela 4: Tempo de processamento para interpolação de 10 mil pontos

Krigagem no mGstat	Tempo
50 dados amostrais:	2,5 minutos
467 dados amostrais:	50 minutos

Os resultados, mostrados na Tabela 4, demonstraram que nesta biblioteca a Krigagem foi implementada da forma tradicional, como descrita no Capítulo 2, onde o aumento dos dados amostrais não é proporcional ao aumento do tempo de processamento.

5 CONCLUSÃO

Este trabalho mostra uma comparação analítica formal entre os métodos da Krigagem e das Redes RBF, utilizados para interpolação espacial. O levantamento bibliográfico feito mostrou que existem trabalhos que mesclam os métodos de Redes Neurais e Krigagem, utilizando-os em conjunto ou utilizando uma característica de um método em outro. Outros trabalhos apenas compararam estatisticamente os métodos.

Verificou-se que a Krigagem possui uma maior base de Geoestatística que as Redes RBF por causa das restrições impostas pelo semivariograma.

Na comparação entre os métodos, feita neste trabalho, foi provado matematicamente que Krigagem e as Redes RBF produzem resultados idênticos quando são utilizadas a mesma função de ativação e semivariograma e com parâmetros equivalentes. Porém, o processo de estimativa dos parâmetros da função de ativação e do semivariograma diverge. Além disso, a equivalência dos parâmetros do semivariograma e da função de ativação das Redes RBF foi demonstrada, e apesar do semivariograma possuir dois parâmetros extras, chamados de *pepita* e *limiar*, eles não influenciam na interpolação. Por causa do diferente processo de estimativa destes parâmetros, os trabalhos que compararam a Krigagem e as Redes RBF concluíram que os métodos são diferentes.

Foi mostrado que a Krigagem e as Redes RBF possuem a mesma estrutura, e por isso podem compartilhar as melhorias feitas em cada método. Por exemplo, as Redes RBF não possuem um método para estimar o intervalo de predição que a Krigagem possui. Outro exemplo é que normalmente as Redes RBF não utilizam um polinômio na sua interpolação como a Krigagem, que melhora o resultado da interpolação. Além disso, toda a base estatística que é usada para realizar a interpolação da Krigagem pode se estender às Redes RBF, devido à equivalência. Porém, a utilização de um procedimento de um método em outro deve ser feita com cautela. Alguns destes procedimentos se baseiam nas premissas de cada método e podem não fazer sentido utilizá-la no outro método sem esse cuidado. Por exemplo, para estimar o erro de interpolação nas Redes RBF é necessário que a função de ativação seja ajustada pelo semivariograma experimental, e que seja um semivariograma válido, de acordo com o exposto na seção da Krigagem.

Os cálculos de complexidade em notação “Big O” demonstraram que as Redes RBF são muito mais eficientes computacionalmente que a Krigagem. Isso se deve ao fato que as Redes RBF só calculam os pesos uma vez e os reutilizam em todos os dados interpolados, enquanto a Krigagem precisa recalcular os pesos para cada ponto a ser interpolado.

Foram feitos experimentos que demonstraram na prática a complexidade dos métodos comparados, onde as Redes RBF obtiveram resultados de mais eficiência computacional quando foi aumentada a quantidade de dados amostrais. Já o aumento de dados interpolados gerou um aumento relativo no tempo de processamento igual nos dois métodos, e esse resultado era esperado pela complexidade calculada, que não inclui a quantidade de dados interpolados.

Foi verificado, também através de experimentos, que a biblioteca mGstat do MATLAB implementa a Krigagem da forma tradicional em que foi descrita nesta dissertação, o que a deixa menos eficiente computacionalmente que as Redes RBF.

Em outro experimento, utilizando o software comercial GS+, foi verificado que este implementa a Krigagem de forma otimizada e não sofre o problema que a implementação original sofre em relação ao perda de eficiência computacional relacionada ao aumento de dados amostrais. Porém, o software limita a quantidade de dados amostrais que podem ser usados, provavelmente para evitar que o processamento se torne muito custoso.

Apesar de o objetivo inicial ter sido melhorar os métodos, ele não foi totalmente atingido. Isso porque, apesar de ser possível uma maior eficiência no método da Krigagem, o seu tempo de processamento será igual ao do método das Redes RBF, que já conhecido. Além disso, não foi obtida nenhuma melhoria em relação a diminuir o erro de interpolação.

5.1 TRABALHOS FUTUROS

Existem algumas áreas que são de interesse para pesquisas futuras, mas que não foram estudadas e analisadas com profundidade durante a elaboração desta dissertação.

Um estudo possível é verificar se a adição de dados interpolados nos dados amostrais melhora a interpolação. Apesar de ficar um algoritmo computacionalmente mais custoso, a intenção é que as áreas sem dados amostrais tenham uma interpolação melhorada.

Outra área que existe muita possibilidade de pesquisa é a de estudar algoritmos que estimem melhor o semivariograma. Afinal, a Krigagem só obtém uma boa interpolação quando o modelo de semivariograma é bem estimado.

As Redes RBF exatas foram analisadas nesta dissertação. Porém, as Redes RBF mais popularmente usadas são as que centralizam as funções de base radial dos neurônios nos agrupamentos de dados amostrais, ou em posições aleatórias e bem distribuídas. Uma linha de pesquisa interessante é

verificar matematicamente se é possível utilizar a variância da interpolação, oriundo da Krigagem, nestes tipos de Redes RBF.

Além disso, existem muitos tipos de Redes RBF e Krigagem não analisados, como tipos em que os parâmetros da função de ativação são diferentes para cada neurônio, ou tipos de Krigagem que utilizam uma janela de dados amostrais para calcular o semivariograma, também gerando semivariogramas com diferentes parâmetros para cada dado a ser interpolado.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKCIN, H.; CELIK, C. T. Performance of Artificial Neural Networks on Kriging Methods in Modeling Local Geoid. *Boletim de Ciências Geodésicas*, v. 19, p. 84–97, 2013.
- BADEL, M.; ANGORANI, S.; PANAHI, M. S. The application of median indicator kriging and neural network in modeling mixed population in an iron ore deposit. *Computers & Geosciences*, Elsevier, v. 37, n. 4, p. 530–540, abr. 2011. ISSN 00983004. <<http://linkinghub.elsevier.com/retrieve/pii/S0098300410003006>>.
- BERRY, J. K. Topic 7 : Linking Data Space and Beware the Slippery Surfaces of GIS Modeling. In: *Beyond Mapping III*. [S.l.]: GeoTec Media Publishers, 2007.
- BERTSEKAS, D. P. *Constrained Optimization and Lagrange Multiplier Methods*. [S.l.]: Academic Press, 1982. 410 p.
- BISHOP, C. M. Neural networks for pattern recognition. In: . [S.l.]: Oxford University, 1996. p. 504.
- BROOMHEAD, D. S.; LOWE, D. Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, v. 2, p. 321–355, 1988.
- CARR, J. C.; FRIGHT, W. R.; BEATSON, R. K. Surface Interpolation with Radial Basis Functions for Medical Imaging. *IEEE Transactions on Medical Imaging*, v. 16, n. 1, p. 96–107, 1997.
- CELLURA, M. et al. Wind speed spatial estimation for energy planning in Sicily: A neural kriging application. *Renewable Energy*, v. 33, n. 6, p. 1251–1266, jun. 2008. ISSN 09601481. <<http://linkinghub.elsevier.com/retrieve/pii/S0960148107002649>>.
- CERESETTI, D. et al. Evaluation of classical spatial-analysis schemes of extreme rainfall. *Natural Hazards and Earth System Science*, v. 12, n. 11, p. 3229–3240, nov. 2012. ISSN 1684-9981. <<http://www.nat-hazards-earth-syst-sci.net/12/3229/2012/>>.
- CHAUVET, P.; GALLI, A. *Universal Kriging*. [S.l.]: Ecole Nationale Supérieure des Mines de Paris, 1982. 94 p.
- CLARK, O.; HARPER, W. V. Estimation and Kriging. In: *Practical Geostatistics 2000*. [S.l.]: Ecosse North America, 2000. cap. 10.

CLIMATEMPO. *Climatempo*. 2013. Acessado em: 06/05/2013. <<http://www.climatempo.com.br/>>.

COSTA, J.; PRONZATO, L.; THIERRY, E. A Comparison Between Kriging and Radial Basis Function Networks for Nonlinear Prediction. *EEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, II, p. 726–730, 1999.

CRESSIE, N. A. C. *Statistics For Spatial Data Revised Edition*. [S.l.]: Wiley-Interscience, 1993. 928 p. ISBN 0-471-00255-0.

DALEZIOS, N. R.; LOUKAS, A.; BAMPZELIS, D. Universal kriging of hail impact energy in Greece. *Physics and Chemistry of the Earth*, v. 27, n. 23-24, p. 1039–1043, jan. 2002. ISSN 14747065. <<http://linkinghub.elsevier.com/retrieve/pii/S1474706502001377>>.

DEMYANOV, V. et al. Neural Network Residual Kriging Application for Climatic Data. *Journal of Geographic Information and Decision Analysis*, v. 2, n. 2, p. 215–232, 1998.

DEMYANOV, V. et al. Wavelet analysis residual kriging vs. neural network residual kriging. *Stochastic Environmental Research and Risk Assessment*, v. 15, n. 1, p. 18–32, mar. 2001. ISSN 14363240. <<http://www.springerlink.com/index/10.1007/s004770000056>>.

DUBOIS, G. Spatial Interpolation Comparison 97:Foreword and Introduction. *Journal of Geographic Information and Decision Analysis*, v. 2, n. 2, p. 1–11, 1998.

GAMMADESIGN. *GS+: Geostatistics for the Environmental Sciences*. 2013. Acessado em: 06/05/2013. <<http://www.gammadesign.com/>>.

GNEITING, T.; SASVÁRI, Z.; SCHLATHER, M. Analogies and Correspondences Between Variograms and Covariance Functions. *NRCSE Technical Report Series*, v. 56, 2000.

GRINGARTEN, E.; DEUTSCH, C. V. Teacher's Aide - Variogram Interpretation and Modeling. *Mathematical Geology*, v. 33, n. 4, p. 507–534, 2001.

HANSEN, T. M. *mGstat : A Geostatistical Matlab toolbox*. 2013. Acessado em: 06/05/2013. <<http://mgstat.sourceforge.net/>>.

HARTIGAN, J. A.; WONG, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of Royal Statistical Society. Series C (Applied Statistics)*, v. 28, n. 1, p. 100–108, 1979.

HAYKIN, S. *Neural Networks and Learning Machines*. [S.l.]: Prentice Hall, 2008. 936 p.

KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. *Science (New York, N.Y.)*, v. 220, n. 4598, p. 671–80, maio 1983. ISSN 0036-8075. <<http://www.ncbi.nlm.nih.gov/pubmed/17813860>>.

KLEIN, D. *Lagrange Multipliers without Permanent Scarring*. 6 p.

KNUTH, D. E. Fundamental Algorithms. In: *The Art of Computer Programming*. [S.l.]: Addison-Wesley, 1968. v. 1, cap. 1.

KOIKE, K.; MATSUDA, S.; GU, B. Evaluation of Interpolation Accuracy of Neural Kriging with Application to Temperature-Distribution Analysis. *Mathematical Geology*, v. 33, n. 4, p. 421–448, 2001.

LI, J.; HEAP, A. D. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Record 200. [S.l.]: Geoscience Australia, 2008. 137 p.

LI, Q.-q. et al. Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach. *Catena*, Elsevier B.V., v. 104, p. 210–218, maio 2013. ISSN 03418162. <<http://linkinghub.elsevier.com/retrieve/pii/S0341816212002494>>.

LIN, G.-F.; CHEN, L.-H. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. *Journal of Hydrology*, v. 288, n. 3-4, p. 288–298, mar. 2004. ISSN 00221694. <<http://linkinghub.elsevier.com/retrieve/pii/S0022169403004244>>.

MARQUARDT, D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, v. 11, n. 2, p. 431–441, 1963.

MATHWORKS. *Matlab: The Language of Technical Computing*. 2013. Acessado em: 06/05/2013. <<http://www.mathworks.com>>.

MATÍAS, J. M. et al. Comparison of Kriging and Neural Networks With Application to the Exploitation of a Slate Mine. *Mathematical Geology*, v. 36, n. 4, p. 463–486, 2004.

MORDECAI, A. *Nonlinear Programming: Analysis and Methods*. [S.l.]: Dover Publications, 2003. 528 p.

OCTAVE. *GNU Octave*. 2013. Acessado em: 06/05/2013.
<<http://www.gnu.org/software/octave/>>.

PRESS, W. H. et al. Modeling of Data. In: *Numerical Recipes in C*. [S.l.]: Cambridge University Press, 1992. cap. 15, p. 656–706. ISBN 0521431085.

SAKATA, S.; ASHIDA, F.; ZAKO, M. Comparative study on gradient and Hessian estimation using the Kriging method and neural network approximation. *Mathematical and Computer Modelling*, Elsevier Ltd, v. 51, n. 3-4, p. 309–319, fev. 2010. ISSN 08957177.
<<http://linkinghub.elsevier.com/retrieve/pii/S0895717709002763>>.

SATHYANARAYANAMURTHY, H.; CHINNAM, R. B. Metamodels for variable importance decomposition with applications to probabilistic engineering design. *Computers & Industrial Engineering*, Elsevier Ltd, v. 57, n. 3, p. 996–1007, out. 2009. ISSN 03608352.
<<http://linkinghub.elsevier.com/retrieve/pii/S0360835209001041>>.

SIL, S.; SRINIVASAN, S. Stochastic simulation of fracture strikes using seismic anisotropy induced velocity anomalies. *Exploration Geophysics*, v. 40, n. 3, p. 257, 2009. ISSN 0812-3985.
<<http://www.publish.csiro.au/?paper=EG08129>>.

SRINIVASAN, B. V.; DURAISWAMI, R.; MURTUGUDDE, R. Efficient kriging for real-time spatio-temporal interpolation Linear kriging. *20th Conference on Probability and Statistics in the Atmospheric Sciences*, p. 1–8, 2011.

STEIN, A.; MEER, F. van der; GORTE, B. Spatial Prediction by Linear Kriging. In: *Spatial Statistics for Remote Sensing*. [S.l.]: Kluwers Academic Publishers, 2002. cap. 6, p. 100–131.

TROCHU, F. A Contouring Program Based on Dual Kriging Interpolation. *Engineering with Computers*, v. 9, p. 160–177, 1993.

VENKATRAM, A. On the use of kriging in the spatial analysis of acid precipitation data. *Atmospheric Environment*, v. 22, n. 9, p. 1963–1975, 1988.

YERRABOLU, P. et al. Correction Model-Based ANN Modeling Approach for the Estimation of Radon Concentrations in Ohio. *Environmental Progress & Sustainable Energy*, v. 00, n. 00, p. 1–11, 2012.

ZHANG, B.; GOVINDARAJU, R. S. Spatial Interpolation of Hydraulic Conductivity using Radial-Basis Function Networks. *Water Resources*, p. 1–10, 2000.

ZHENG, Z. et al. Spatial estimation of soil moisture and salinity with Neural Kriging. *Computer and Computing Technologies in Agriculture II*, v. 2, p. 1227–1237, 2009.