

4

Experimentos e comparações

É muito subjetivo escolher as características que serão usadas para comparar vários métodos e como elas podem ser utilizadas para se fazer uma análise final. Dentre essas características as mais utilizadas para comparar métodos de interpolação são geralmente: aspectos visuais, precisão, tempo de execução, sensibilidade aos parâmetros e facilidade de implementação.

Para analisar essas características foram feitos três experimentos para testar o novo método, são eles:

1. Reconstrução de uma função real de uma variável que é dada por uma equação,
2. Reconstrução de uma função real de duas variáveis que é dada por um grid bidimensional, e
3. Reconstrução de um campo de vetores no plano bidimensional.

Para cada um desses três experimentos foram analisados os seguintes núcleos:

- Núcleo gaussiano: $k_{gauss}(\mathbf{x}, \mathbf{z}) = \exp(-\frac{(\|\mathbf{x} - \mathbf{z}\|)^2}{\theta})$,
- Núcleo wave: $k_{wave}(\mathbf{x}, \mathbf{z}) = \begin{cases} \frac{\theta}{\|\mathbf{x} - \mathbf{z}\|} \sin(\frac{\|\mathbf{x} - \mathbf{z}\|}{\theta}), & \text{se } \|\mathbf{x} - \mathbf{z}\| \neq 0 \\ 1, & \text{caso contrário} \end{cases}$,
- Núcleo wavelets: $k_{wvls}(\mathbf{x}, \mathbf{z}) = \prod_{i=1}^n \cos(1.75 \frac{(x_i - z_i)}{\theta}) \exp(-\frac{\|x_i - z_i\|^2}{2\theta^2})$,
- Núcleo polinomial: $k_{pol}(\mathbf{x}, \mathbf{z}) = (\langle \mathbf{x}, \mathbf{z} \rangle)^\theta$,
- Núcleo exponencial: $k_{exp}(\mathbf{x}, \mathbf{z}) = \exp(-\frac{\|\mathbf{x} - \mathbf{z}\|}{\theta})$,
- Núcleo circular: $k_{circ}(\mathbf{x}, \mathbf{z}) = \begin{cases} \frac{2}{\pi} \arccos(\frac{\|\mathbf{x} - \mathbf{z}\|}{\theta}) - \frac{2}{\pi} \frac{\|\mathbf{x} - \mathbf{z}\|}{\theta} \sqrt{1 - \frac{\|\mathbf{x} - \mathbf{z}\|^2}{\theta^2}}, & \text{se } \|\mathbf{x} - \mathbf{z}\| < \theta \\ 0, & \text{caso contrário} \end{cases}$,
- Núcleo racional quadrático: $k_{RQ}(\mathbf{x}, \mathbf{z}) = 1 - \frac{\|\mathbf{x} - \mathbf{z}\|^2}{\|\mathbf{x} - \mathbf{z}\|^2 + \theta}$.

Reparem que todos os núcleos acima são parametrizados por um único número θ , e o parâmetro p do método de Shepard foi escolhido como sendo igual a dois em todos os experimentos.

O objetivo aqui é comparar os resultados do método de Shepard baseado em núcleos com o método original, o que corresponde a usar o núcleo euclidiano $k_{euc}(\mathbf{x}, \mathbf{z}) = \langle \mathbf{x}, \mathbf{z} \rangle$ no cálculo da distância d_k . Assim o método de Shepard original vira um caso particular do método proposto.

Sobre a facilidade de implementação, a modificação proposta só requer uma substituição de uma chamada do procedimento de interpolação, que troca o cálculo da distância euclidiana pela distância usando o núcleo dado.

Todos os testes aqui apresentados foram implementados na plataforma R (software livre disponível no sítio: <http://www.r-project.org/>).

4.1

Reconstrução de uma função real de uma variável

Nessa primeira análise da modificação do método de interpolação Shepard será utilizada a seguinte função real de uma variável:

$$f(x) = 15e^{\frac{-1}{1-(2x-1)^2}} \left(\frac{3}{4}e^{-\frac{(9x-2)^2}{4}} + \frac{3}{4}e^{-\frac{(9x+1)^2}{49}} + \frac{1}{2}e^{-\frac{(9x-7)^2}{4}} - \frac{1}{5}e^{-(9x-2)^2} \right)$$

Para fazer as comparações de uma forma mais robusta estatisticamente, a função f foi amostrada em 100 pontos igualmente espaçados no intervalo $[0, 1]$, depois a partir desse conjunto de 100 pontos foram construídos de forma aleatória 10 conjuntos de 10 pontos, 10 conjuntos de 20 pontos e 10 conjuntos de 30 pontos.

Para fazer a análise da precisão foi feito o seguinte procedimento. Para cada um dos 10 conjuntos de N pontos ($N \in \{10, 20, 30\}$), foram calculadas as estatísticas (mínimo, máximo, média, desvio padrão, mediana e máximo desvio absoluto em relação a mediana (MAD)) dos erros quadráticos $(\hat{f}_k(x_i) - f(x_i))^2$ nos 100 pontos igualmente espaçados no intervalo $[0, 1]$. Depois disso, foram recalculadas as mesmas estatísticas dessas estatísticas de erros nos 10 conjuntos.

As tabelas 4.1, 4.2 e 4.3 mostram os resultados para conjuntos de 10, 20 e 30 pontos, respectivamente. Em todas elas, os números em negrito representam os núcleos que tiveram um desempenho melhor que o núcleo euclidiano, ou seja melhor que o método de Shepard original. No teste com 10 pontos, chama a atenção o fato que nas estatísticas média e mediana todos os núcleos tiveram desempenho melhor que o núcleo euclidiano. No teste com 20 pontos, o que ficou marcante foi o fato de que o núcleo polinomial teve um excelente desempenho, vale lembrar que dessa lista ele é o único núcleo que é não estacionário. Por outro lado, para os testes com 30 pontos os núcleos gaussiano e racional quadrático (que são estacionários e isotrópicos) tiveram o

melhor resultado.

Tabela 4.1: Tabela comparativa dos erros quadráticos considerando interpolações com 10 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,20257	0,13589	0,13603	0,13567	0,13603	0,11490	0,11371	0,13491
Máximo	1,46871	1,55226	1,55160	1,55327	1,55156	1,74596	1,74504	1,55676
Média	0,55544	0,54601	0,54586	0,54616	0,54585	0,67505	0,67210	0,54716
D.Padrão	0,38787	0,43763	0,43733	0,43811	0,43732	0,53985	0,53941	0,43976
Mediana	0,55016	0,41567	0,41622	0,41479	0,41625	0,48007	0,47652	0,41173

Tabela 4.2: Tabela comparativa dos erros quadráticos considerando interpolações com 20 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,07940	0,06056	0,06072	0,06068	0,07783	0,08718	0,08057	0,06046
Máximo	2,54691	2,63017	2,69071	2,67497	2,48557	2,90684	2,92058	2,64277
Média	0,81114	0,77114	0,79872	0,79196	0,82056	0,81474	0,81979	0,76861
D.Padrão	0,85064	0,83312	0,86550	0,85721	0,83066	0,87129	0,88308	0,83319
Mediana	0,55243	0,57273	0,57440	0,57430	0,53726	0,67344	0,67082	0,57395
MAD	0,58213	0,62785	0,62915	0,62936	0,55209	0,65224	0,66460	0,63021

Tabela 4.3: Tabela comparativa dos erros quadráticos considerando interpolações com 30 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,06604	0,05911	0,05878	0,05885	0,06621	0,08637	0,08508	0,05932
Máximo	0,87803	0,84492	0,89364	0,883679	0,92309	1,09352	1,07400	0,83987
Média	0,31840	0,31576	0,31991	0,319121	0,35634	0,38067	0,37601	0,31595
D.Padrão	0,29143	0,30323	0,31147	0,309833	0,30186	0,34446	0,34687	0,30337
Mediana	0,19421	0,16435	0,16725	0,16650	0,25637	0,25975	0,24890	0,16276
MAD	0,13346	0,11111	0,10944	0,10985	0,22196	0,22914	0,22378	0,11216

Em todas essas comparações, ficou claro que a modificação proposta pode gerar resultados melhores que o método de Shepard original quando é feita uma boa escolha do parâmetro θ . Na tabela 4.4 estão listados os valores de θ utilizados na comparação. Para cada núcleo k e para cada valor de $N \in \{10, 20, 30\}$, o valor de θ “ótimo” foi determinado escolhendo num conjunto pré-determinado de valores e aquele que produzia o menor erro médio quadrático para um dos 10 conjuntos com N amostras.

Tabela 4.4: Tabela com os valores de θ nas reconstruções com 10, 20, e 30 pontos.

Valores de θ	10 pontos	20 pontos	30 pontos
k_{gauss}	2	0.4	0.4
k_{wave}	4	0.4	0.4
k_{wlts}	2	1	1
k_{pol}	1	2	1.6
k_{exp}	10	2	2
k_{circ}	π	π	π
k_{RQ}	2	0.4	0.4

Para comparar o aspecto visual das reconstruções por diferentes núcleos foi escolhido um conjunto com 10 pontos. As figuras 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 e 4.8 ilustram o gráfico da função interpoladora e o gráfico da função f original para os núcleos euclidiano, gaussiano, wave, wavelets, polinomial, exponencial, circular, e racional quadrático. Nesse experimento de reconstrução de uma função real de uma só variável, existem poucas diferenças visuais entre eles. Vale notar que os núcleos circular e exponencial tiveram os gráficos visualmente mais parecidos com a função original para esse conjunto, mas na média considerando os 10 conjuntos de 10 pontos eles tiveram desempenho pior.

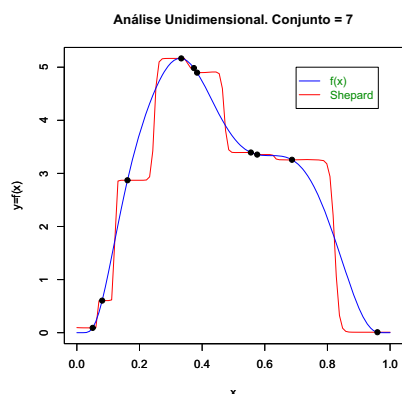


Figura 4.1: Função original e a interpolação de Shepard usando o núcleo euclidiano.

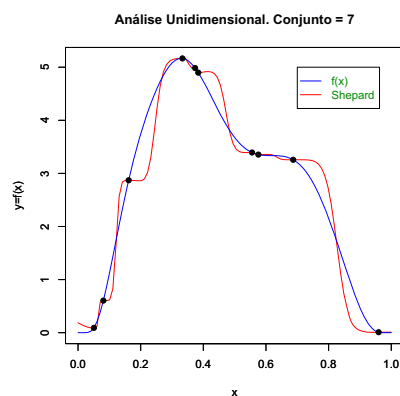


Figura 4.2: Função original e interpolação de Shepard usando o núcleo gaussiano parâmetro $\theta = 2$.

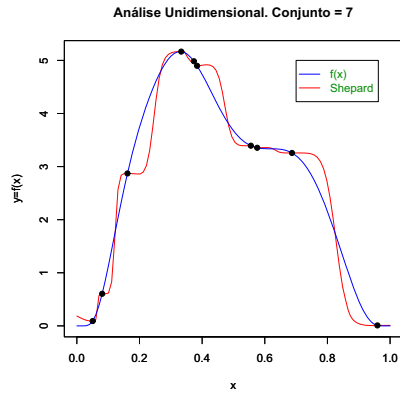


Figura 4.3: Função original e interpolação de Shepard usando o núcleo wave parâmetro $\theta = 4$.

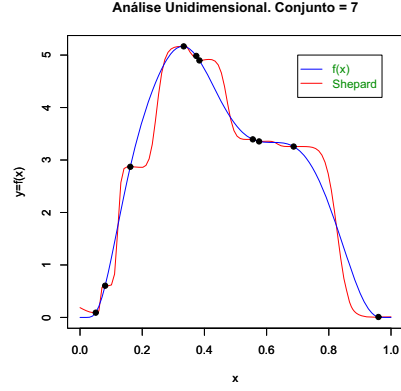


Figura 4.4: Função original e interpolação de Shepard usando o núcleo wavelets parâmetro $\theta = 2$.

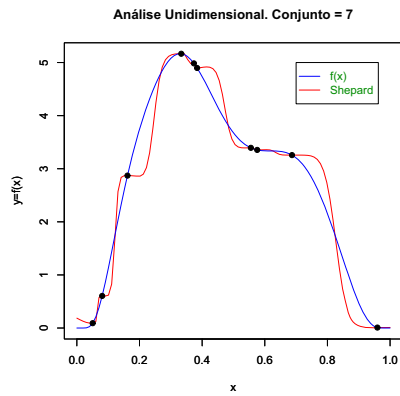


Figura 4.5: Função original e interpolação de Shepard usando o núcleo polinomial parâmetro $\theta = 1$.

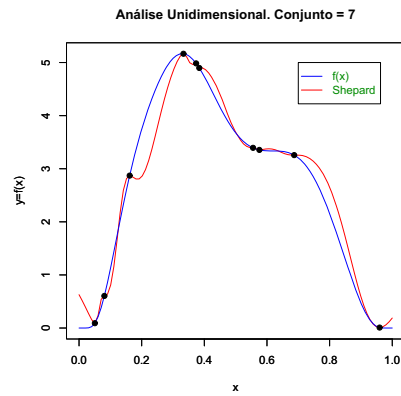


Figura 4.6: Função original e interpolação de Shepard usando o núcleo exponencial parâmetro $\theta = 10$.

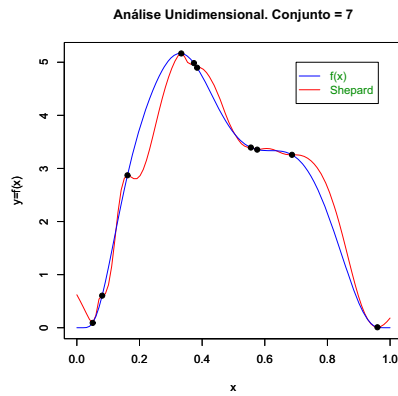


Figura 4.7: Função original e interpolação de Shepard usando o núcleo circular parâmetro $\theta = \pi$.

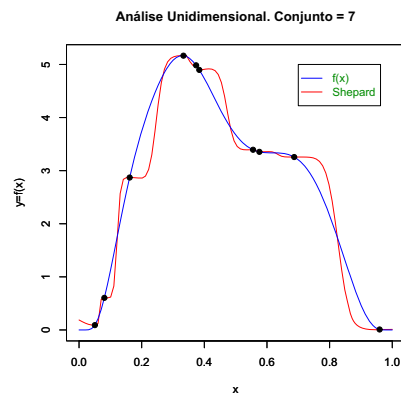


Figura 4.8: Função original e interpolação de Shepard usando o núcleo RQ parâmetro $\theta = 2$.

Para avaliar o tempo de execução foi escolhido comparar o tempo de avaliação da função \hat{f}_k no ponto $x = 0.5$ para cada um dos núcleos em um conjunto de entrada com $N = 30$ pontos. A tabela 4.5 mostra o resultado da comparação.

Tabela 4.5: Tempo de execução para o cálculo de $\hat{f}_k(0.5)$ considerando um conjunto com $N = 30$ pontos.

Núcleo	Tempo	Razão
k_{euc}	0,02s	1,00
k_{gauss}	0,02s	1,00
k_{wave}	0,03s	1,50
k_{wvltts}	0,03s	1,50
k_{pol}	0,02s	1,00
k_{exp}	0,02s	1,00
k_{circ}	0,04s	2,0
k_{RQ}	0,02s	1,00

4.2

Reconstrução de uma função real de duas variáveis

Esse segundo teste utilizará o dado denominado `volcano`, disponível na própria plataforma R, que fornece os dados topográficos de um vulcão conhecido como Maunga Whau. Esse dado está representado por um matriz de 87 linhas por 61 colunas, totalizando portanto, 5307 valores que correspondem a altitude do terreno. O domínio desse dado, foi assumido ser o retângulo $[0, 1] \times [0, \frac{87}{61}]$. Assim, do ponto de vista geométrico esse dado pode ser interpretado como um

grid em que cada uma de suas células tem área $\frac{1}{61} \times \frac{1}{61}$, e cada vértice tem informação sobre a elevação topográfica no ponto onde o vértice está localizado.

Para fazer as comparações foram selecionados de forma aleatória, assumindo uma distribuição uniforme, 100 pontos dos 5307 dados do `volcano`. Após esta escolha aleatória, desses 100 pontos foram sorteados 10 conjuntos com 20 pontos, 10 conjuntos de 30 pontos, 10 conjuntos de 40 pontos e 10 conjuntos de 50 pontos, para todos esses sorteios foi assumida também uma distribuição uniforme.

A análise feita para esse caso é bastante similar a do caso anterior. Para cada um dos 10 conjuntos de N pontos ($N \in \{20, 30, 40, 50\}$), foram calculadas as estatísticas (mínimo, máximo, média, desvio padrão, mediana e máximo desvio absoluto em relação a mediana (MAD)) dos erros quadráticos $(\hat{f}_k(\mathbf{x}_i) - z(\mathbf{x}_i))^2$ nos 100 pontos inicialmente sorteados, onde $z(\mathbf{x}_i)$ é o valor da elevação topográfica do ponto \mathbf{x}_i , que está disponível na matriz `volcano`. Depois disso, foram recalculadas as mesmas estatísticas dessas estatísticas de erros nos 10 conjuntos.

As tabelas 4.6, 4.7, 4.8 e 4.9 mostram os resultados para conjuntos de 20, 30, 40 e 50 pontos, respectivamente. Em todas elas, os números em negrito representam os núcleos que tiveram um desempenho melhor que o núcleo euclidiano, ou seja, melhor que o método de Shepard original. Vale destacar nesse contexto bidimensional que o núcleo racional quadrático obteve um excelente desempenho para todos os conjuntos de N testados em todas as estatísticas calculadas. O núcleo gaussiano teve também um grande desempenho. No caso bidimensional, o fato de que um determinado núcleo é estacionário e isotrópico pode fazer diferença numérica e visual, como foi o caso.

Tabela 4.6: Tabela comparativa dos erros quadráticos considerando interpolações com 20 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,31959	0,29384	0,30707	0,29976	0,30793	0,32978	0,31971	0,29692
Máximo	0,66867	0,63362	0,70970	0,66608	0,71546	0,66277	0,66753	0,64102
Média	0,47029	0,43531	0,47362	0,45113	0,47524	0,47642	0,47014	0,44033
D.Padrão	0,11615	0,11640	0,13306	0,11986	0,13436	0,11576	0,11613	0,11562
Mediana	0,44808	0,40330	0,43560	0,41356	0,43660	0,45298	0,44725	0,40870
MAD	0,11793	0,10616	0,11453	0,09285	0,11538	0,12023	0,11703	0,09666

Tabela 4.7: Tabela comparativa dos erros quadráticos considerando interpolações com 30 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,28076	0,23505	0,23135	0,29976	0,30017	0,28814	0,28072	0,23953
Máximo	0,72255	0,65661	0,73865	0,66608	0,73357	0,73097	0,72248	0,66821
Média	0,44867	0,39521	0,42377	0,45113	0,48306	0,45522	0,44844	0,40331
D.Padrão	0,14293	0,13709	0,16578	0,11986	0,14538	0,13879	0,14242	0,14001
Mediana	0,42534	0,37727	0,39106	0,41356	0,48511	0,42842	0,42487	0,38557
MAD	0,12688	0,11985	0,13932	0,09285	0,13672	0,12361	0,12647	0,11966

Tabela 4.8: Tabela comparativa dos erros quadráticos considerando interpolações com 40 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,28076	0,23505	0,23135	0,29976	0,30017	0,28814	0,28072	0,23953
Máximo	0,72255	0,65661	0,73865	0,66608	0,73357	0,73097	0,72248	0,66821
Média	0,44867	0,39521	0,42377	0,45113	0,48306	0,45522	0,44844	0,40331
D.Padrão	0,14293	0,13709	0,16578	0,11986	0,14538	0,13879	0,14242	0,14001
Mediana	0,42534	0,37727	0,39106	0,41356	0,48511	0,42842	0,42487	0,38557
MAD	0,12688	0,11985	0,13932	0,09285	0,13672	0,12361	0,12647	0,11966

Tabela 4.9: Tabela comparativa dos erros quadráticos considerando interpolações com 50 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,21402	0,17560	0,18294	0,17978	0,18287	0,21917	0,21442	0,17820
Máximo	0,45116	0,40126	0,45738	0,42504	0,45937	0,45489	0,45114	0,40574
Média	0,31954	0,27352	0,28277	0,27757	0,28315	0,32469	0,31990	0,27702
D.Padrão	0,07753	0,07273	0,08686	0,07963	0,08730	0,07678	0,07734	0,07368
Mediana	0,31332	0,26415	0,26425	0,26344	0,26439	0,31774	0,31361	0,26847
MAD	0,06460	0,06400	0,07028	0,07372	0,07005	0,0659	0,06452	0,06445

Em todas essas comparações, ficou claro novamente que a modificação proposta pode gerar resultados melhores para uma boa escolha do parâmetro θ . Na tabela 4.10 estão listados os valores de θ utilizados na comparação. Para cada núcleo k e para cada valor de $N \in \{20, 30, 40, 50\}$, o valor de θ “ótimo” foi determinado escolhendo num conjunto predeterminado de valores e aquele que produzia o menor erro médio quadrático para um dos 10 conjuntos com N amostras.

Tabela 4.10: Tabela com os valores de θ nas reconstruções com 20, 30, 40 e 50 pontos.

	20 Pontos	30 Pontos	40 Pontos	50 Pontos
k_{gauss}	0,4	0,4	0,4	0,4
k_{wave}	1	1	1	1
k_{wlts}	1	1	1	1
k_{pol}	2	4	2	2
k_{exp}	2	2	2	2
k_{circ}	π	π	π	π
k_{RQ}	0,4	0,4	0,4	0,4

Para comparar o aspecto visual das reconstruções por diferentes núcleos foi escolhido um conjunto com 50 pontos . As figuras 4.9 e 4.18 ilustram a superfície original do dado **volcano**. As figuras (4.10 e 4.19), (4.11 e 4.20), (4.12 e 4.21), (4.13 e 4.22), (4.14 e 4.23), (4.15 e 4.24), (4.16 e 4.25) e (4.17 e 4.26) ilustram o gráfico da função interpoladora para os núcleos euclidiano, gaussiano, wave, wavelets, polinomial, exponencial, circular, e racional quadrático. Nesse experimento de reconstrução de uma função real de duas variáveis, existem diferenças visuais entre eles, por exemplo o gráfico do núcleo gaussiano (4.11 e 4.20) está bem diferente do gráfico do núcleo circular (4.16 e 4.25).

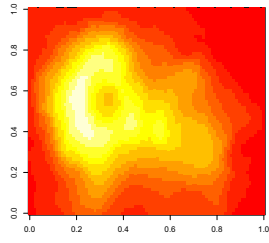


Figura 4.9: Imagem original gerada pelo programa estatístico R.

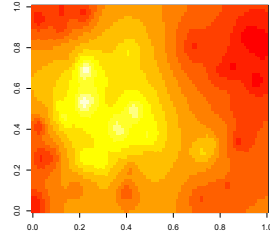


Figura 4.10: Imagem núcleo euclidiano amostra 50 pontos.

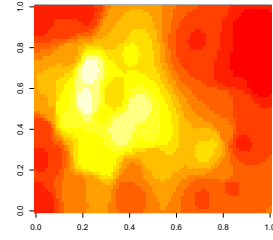


Figura 4.11: Imagem núcleo gaussiano parâmetro $\theta = 0.4$.

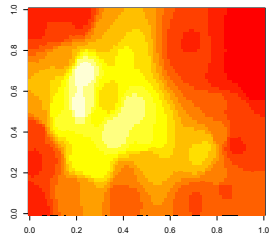


Figura 4.12: Imagem núcleo wave parâmetro $\theta = 1$.

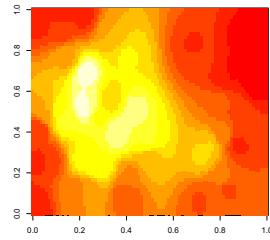


Figura 4.13: Imagem núcleo wavelets parâmetro $\theta = 1$.

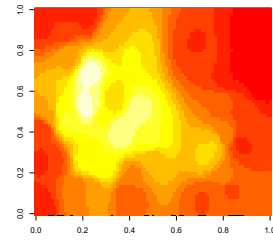


Figura 4.14: Imagem núcleo polinomial parâmetro $\theta = 1$.

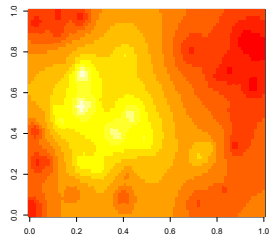


Figura 4.15: Imagem núcleo exponencial parâmetro $\theta = 2$.

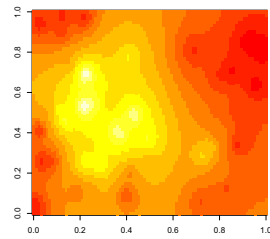


Figura 4.16: Imagem núcleo circular parâmetro $\theta = \pi$.

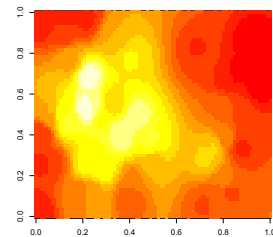


Figura 4.17: Imagem núcleo racional quadrático parâmetro $\theta = 0.4$.

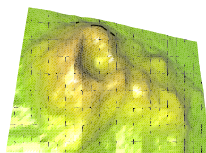


Figura 4.18:
Superfície original
do vulcão gerada pelo
programa estatístico
R.

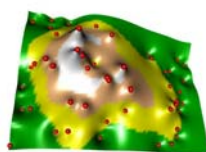


Figura 4.19:
Superfície núcleo
euclidiano com
amostra 50 pontos.

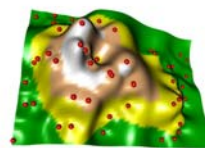


Figura 4.20:
Superfície núcleo
gaussiano parâmetro
 $\theta = 0.4$.

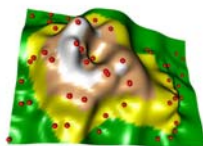


Figura 4.21: Su-
perfície núcleo wave
parâmetro $\theta = 1$.

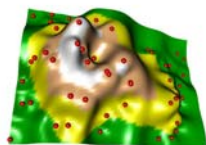


Figura 4.22: Su-
perfície núcleo
wavelets parâmetro
 $\theta = 1$.

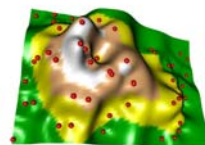


Figura 4.23:
Superfície núcleo
polinomial parâmetro
 $\theta = 1$.

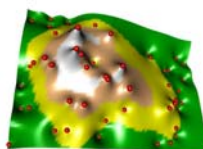


Figura 4.24:
Superfície núcleo ex-
ponencial parâmetro
 $\theta = 2$.

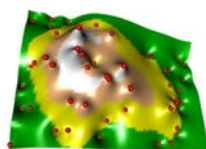


Figura 4.25:
Superfície núcleo
circular parâmetro
 $\theta = \pi$.

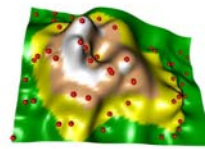


Figura 4.26:
Superfície núcleo
racional quadrático
parâmetro $\theta = 0.4$.

Para exemplificar a convergência do erro para um determinado núcleo, por exemplo, o núcleo gaussiano para um parâmetro $\theta = 0.4$, foram feitos dois gráficos. A figura 4.27 mostra o gráfico da média com um intervalo de mais ou menos o desvio padrão dos erros quadrados médios nos 10 conjuntos para $N \in \{20, 30, 40, 50\}$. Já a figura 4.28 mostra o gráfico da mediana com um intervalo de mais ou menos o MAD dos erros quadrados médios nos 10 conjuntos para $N \in \{20, 30, 40, 50\}$. E a tabela 4.11 mostra todas as estatísticas. Como esperado, observa-se que o erro diminui à medida que a quantidade de pontos aumenta.

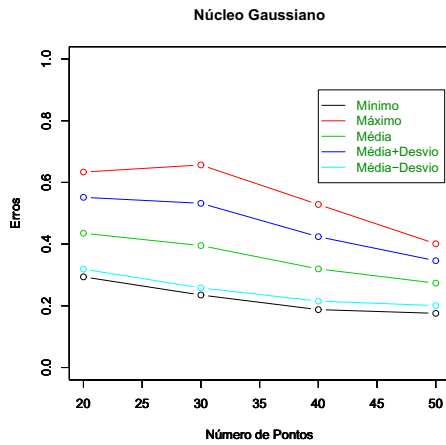


Figura 4.27: Gráfico para análise de convergência para o núcleo gaussiano: média \pm desvio padrão.

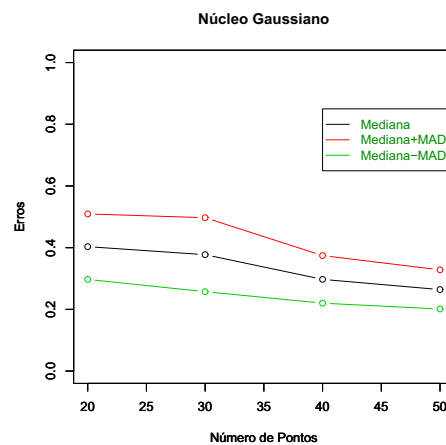


Figura 4.28: Gráfico para análise de convergência para o núcleo gaussiano: mediana \pm MAD.

Tabela 4.11: Análise de convergência para o núcleo gaussiano.

	20 Pontos	30 Pontos	40 Pontos	50 Pontos
Mínimo do Erro Quadrático Médio	0,29384	0,23505	0,18772	0,17560
Máximo do Erro Quadrático Médio	0,63362	0,65661	0,52847	0,40126
Média do Erro Quadrático Médio	0,43531	0,39521	0,3168	0,27352
Desvio Padrão do Erro Quadrático Médio	0,11640	0,13709	0,10440	0,07273
Mediana do Erro Quadrático Médio	0,40330	0,37727	0,29725	0,26415
MAD do Erro Quadrático Médio	0,10616	0,11985	0,07716	0,06400

Para avaliar o tempo de execução foi escolhido comparar o tempo de avaliação da função \hat{f}_k no ponto $\mathbf{x} = (0.5, 0.5)$ para cada um dos núcleos em um conjunto de entrada com $N = 40$ pontos. A tabela 4.12 mostra o resultado da comparação.

Tabela 4.12: Tempo de execução para o cálculo de $\hat{f}_k(0.5, 0.5)$ considerando um conjunto com $N = 40$ pontos.

Núcleo	Tempo	Razão
k_{euc}	0,03s	1,00
k_{gauss}	0,04s	1,33
k_{wave}	0,05s	1,66
k_{wvltts}	0,04s	1,33
k_{pol}	0,04s	1,33
k_{exp}	0,03s	1,00
k_{circ}	0,05s	1,66
k_{RQ}	0,03s	1,00

4.3

Reconstrução de um campo de vetores no plano

O terceiro e último teste utilizará a reconstrução de um campo vetorial no plano. Os vetores desse campo foram obtidos por um dispositivo de PIV (*Particle Image Velocimetry*). Este dado está representado por uma lista de 1000 elementos, onde cada elemento i contém a posição do vetor $\mathbf{x}_i \in \mathbb{R}^2$ e o vetor do campo associado a ele, denotado por $\mathbf{f}_i \in \mathbb{R}^2$. O domínio desse dado, foi assumido ser o retângulo $[0, 1] \times [0, 1]$. Para reconstruir o campo utilizando o método de Shepard proposto, foram reconstruídas cada uma das duas coordenadas do campo separadamente, ou seja, fez-se uma reconstrução usando como dados de entrada, os vetores \mathbf{x}_i associados a primeira coordenada de \mathbf{f}_i , e depois os mesmos vetores \mathbf{x}_i associados a segunda coordenada de \mathbf{f}_i . A função de interpolação obtida será denotada por $\hat{\mathbf{f}} : [0, 1] \times [0, 1] \rightarrow \mathbb{R}^2$.

Para fazer as comparações foram selecionados de forma aleatória, assumindo uma distribuição uniforme sobre os 1000, 10 conjuntos com 10 pontos, 10 conjuntos de 20 pontos, 10 conjuntos de 30 pontos e 10 conjuntos de 40 pontos. A análise feita para esse caso é bastante similar a do caso anterior. Para cada um dos 10 conjuntos de N pontos ($N \in \{20, 30, 40, 50\}$), foram calculadas as estatísticas (mínimo, máximo, média, desvio padrão, mediana e máximo desvio absoluto em relação a mediana (MAD)) dos erros quadrado da diferença da norma (4-1), quadrado da diferença do cosseno do ângulo (4-2) e quadrado da diferença da estimação (4-3).

$${}_1\varepsilon_i^2 = (\|\hat{\mathbf{f}}(\mathbf{x}_i)\| - \|\mathbf{f}_i\|)^2. \quad (4-1)$$

$${}_2\varepsilon_i^2 = \left(\frac{\langle \hat{\mathbf{f}}(\mathbf{x}_i), \mathbf{f}_i \rangle}{\|\hat{\mathbf{f}}(\mathbf{x}_i)\| \cdot \|\mathbf{f}_i\|} \right)^2. \quad (4-2)$$

$${}_3\varepsilon_i^2 = \|\hat{\mathbf{f}}(\mathbf{x}_i) - \mathbf{f}_i\|^2. \quad (4-3)$$

As tabelas 4.13, 4.14, 4.15, 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, e 4.24 mostram os resultados das três diferentes medidas de erros para conjuntos de 10, 20, 30 e 40 pontos, respectivamente. Em relação ao erro dado pela diferença da norma (eq. 4-1), os núcleos exponencial e circular foram os melhores em todas as estatísticas considerando todos os valores de N testados. Esse desempenho se repetiu nas estatísticas de erro considerando a diferença da estimação (eq. 4-3). Por outro lado, ao considerar a diferença do cosseno do ângulo (eq. 4-2) esses mesmos núcleos não tiveram os melhores resultados, que foi obtido, sob um ponto de vista estatístico, pelo núcleo gaussiano.

Tabela 4.13: Estatística dos erros da diferença da norma sobre os conjuntos de 10 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,70212	0,73118	0,73569	0,73504	0,74579	0,70084	0,70203	0,72763
Máximo	0,91904	0,96412	1,03685	0,99051	1,03456	0,90261	0,91750	0,96258
Média	0,78810	0,81603	0,86817	0,84030	0,87458	0,77311	0,78638	0,81642
D.Padrão	0,06834	0,07161	0,09218	0,07949	0,08890	0,06524	0,06790	0,07242
Mediana	0,77161	0,80710	0,85448	0,82827	0,87171	0,75652	0,76996	0,80519
MAD	0,05258	0,06874	0,08124	0,07813	0,06268	0,05033	0,05209	0,06710

Tabela 4.14: Estatística dos erros da diferença do cosseno do ângulo sobre os conjuntos de 10 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,70968	0,70924	0,70754	0,71007	0,69724	0,71114	0,70987	0,70871
Máximo	0,83385	0,82962	0,82576	0,82791	0,82667	0,83493	0,83397	0,82982
Média	0,78958	0,78577	0,77616	0,78047	0,77804	0,79251	0,78989	0,78569
D.Padrão	0,04300	0,04010	0,04117	0,03959	0,04087	0,04260	0,04290	0,04083
Mediana	0,80170	0,79309	0,78588	0,78759	0,78366	0,80363	0,80181	0,79408
MAD	0,03218	0,03807	0,03869	0,03824	0,04190	0,03222	0,03210	0,03665

Tabela 4.15: Estatística dos erros da diferença da estimação sobre os conjuntos de 10 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,91071	0,92893	0,95956	0,94552	0,96486	0,90119	0,90970	0,92925
Máximo	1,33906	1,37458	1,54851	1,43177	1,43177	1,28897	1,33268	1,38485
Média	1,01857	1,04560	1,12129	1,08025	1,12238	0,99693	1,01602	1,04836
D.Padrão	0,12752	0,13427	0,18196	0,14997	0,14538	0,11460	0,12577	0,13694
Mediana	0,99009	1,02134	1,06913	1,04951	1,08715	0,97316	0,98851	1,02073
MAD	0,06527	0,06693	0,10721	0,08021	0,10124	0,06415	0,06540	0,06655

Tabela 4.16: Estatística dos erros da diferença da norma sobre os conjuntos de 20 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,72938	0,76950	0,79789	0,78774	0,80119	0,72025	0,72855	0,76772
Máximo	1,18899	1,37616	1,51241	1,45678	1,61827	1,14611	1,18571	1,36160
Média	0,88000	0,94294	1,20009	0,98340	1,04730	0,85636	0,87769	0,94081
D.Padrão	0,15997	0,19148	0,25985	0,22445	0,28960	0,14200	0,15789	0,19157
Mediana	0,83144	0,87842	0,89186	0,88534	0,89832	0,81151	0,83042	0,87525
MAD	0,11019	0,12180	0,12845	0,12742	0,12481	0,09627	0,11004	0,12266

Tabela 4.17: Estatística dos erros da diferença do cosseno do ângulo sobre os conjuntos de 20 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,77356	0,76677	0,75153	0,75960	0,74651	0,77783	0,77404	0,76612
Máximo	0,83858	0,83682	0,83428	0,83457	0,83390	0,83962	0,83865	0,83641
Média	0,81673	0,81061	0,80014	0,80518	0,79673	0,81958	0,81704	0,81029
D.Padrão	0,02118	0,02303	0,02661	0,02459	0,02774	0,02004	0,02107	0,02317
Mediana	0,82388	0,81663	0,80550	0,81145	0,80122	0,82556	0,82409	0,81672
MAD	0,01601	0,02275	0,02603	0,02473	0,02842	0,01563	0,01596	0,02265

Tabela 4.18: Estatística dos erros da diferença da estimação sobre os conjuntos de 20 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,89774	0,93226	0,96196	0,95278	0,96736	0,88692	0,89680	0,93192
Máximo	1,48217	1,67971	1,83678	1,76891	1,93013	1,43371	1,47830	1,66581
Média	1,07187	1,13900	1,23456	1,18875	1,26834	1,04339	1,06899	1,13805
D.Padrão	0,19445	0,23486	0,30587	0,26935	0,33616	0,17421	0,19224	0,23434
Mediana	1,07187	1,05423	1,11862	1,08935	1,14100	0,98222	0,99862	1,05260
MAD	0,10849	0,14327	0,19584	0,17224	0,21129	0,09205	0,10646	0,14217

Tabela 4.19: Estatística dos erros da diferença da norma sobre os conjuntos de 30 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,70451	0,75126	0,76721	0,76058	0,77351	0,69848	0,70410	0,74775
Máximo	1,05143	1,29878	1,46324	1,38490	1,52285	1,01298	1,04851	1,28413
Média	0,85444	0,97874	1,05606	1,02163	1,08076	0,83109	0,85256	0,97172
D.Padrão	0,11614	0,18465	0,23351	0,21042	0,25079	0,10337	0,11508	0,18093
Mediana	0,84696	0,96370	1,05012	1,01298	1,07481	0,82440	0,84532	0,95531
MAD	0,09557	0,18355	0,21005	0,20061	0,21581	0,09095	0,09492	0,18089

Tabela 4.20: Estatística dos erros da diferença do cosseno do ângulo sobre os conjuntos de 30 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,77062	0,75847	0,74794	0,75236	0,74526	0,77104	0,77063	0,75958
Máximo	0,83518	0,82647	0,82518	0,82510	0,82551	0,83706	0,83534	0,82709
Média	0,81047	0,79917	0,79107	0,79427	0,78853	0,81300	0,81069	0,79998
D.Padrão	0,02015	0,02101	0,02347	0,02244	0,02415	0,02026	0,02150	0,02089
Mediana	0,81139	0,80021	0,79231	0,79584	0,78895	0,81427	0,81159	0,80106
MAD	0,01598	0,01668	0,02042	0,01923	0,02293	0,01515	0,01556	0,01772

Tabela 4.21: Estatística dos erros da diferença da estimação sobre os conjuntos de 30 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,91304	0,97630	1,00021	0,99138	1,00669	0,90322	0,91223	0,97270
Máximo	1,28812	1,56946	1,76688	1,67464	1,83857	1,23114	1,28344	1,55305
Média	1,06242	1,20466	1,30105	1,25892	1,33166	1,30335	1,06002	1,19684
D.Padrão	0,12444	0,20112	0,26341	0,23393	0,28511	0,10877	0,12310	0,19668
Mediana	1,04427	1,18899	1,28772	1,24329	1,32150	1,01560	1,04166	1,18082
MAD	0,12587	0,21471	0,29272	0,25303	0,31653	0,10281	0,12415	0,21038

Tabela 4.22: Estatística dos erros da diferença da norma sobre os conjuntos de 40 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,70839	0,76218	0,77905	0,77333	0,78198	0,69937	0,70773	0,75997
Máximo	1,03733	1,23748	1,48384	1,35778	1,59117	0,97294	1,03047	1,23373
Média	0,79611	0,89315	0,95205	0,92575	0,97248	0,77825	0,79455	0,88798
D.Padrão	0,09926	0,15227	0,22140	0,18527	0,25290	0,08221	0,09732	0,15086
Mediana	0,76737	0,83506	0,86740	0,85361	0,87895	0,75504	0,76636	0,83147
MAD	0,04756	0,07190	0,08079	0,07729	0,08190	0,04435	0,04701	0,06919

Tabela 4.23: Estatística dos erros da diferença do cosseno do ângulo sobre os conjuntos de 40 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,78796	0,78085	0,76772	0,77427	0,76368	0,79064	0,78825	0,78034
Máximo	0,84315	0,83391	0,82804	0,83083	0,82649	0,84517	0,84332	0,83434
Média	0,81377	0,80148	0,79156	0,79584	0,78867	0,81693	0,81408	0,80179
D.Padrão	0,01731	0,01676	0,01815	0,01747	0,01858	0,01715	0,01728	0,01689
Mediana	0,81173	0,79900	0,79010	0,79390	0,78699	0,81529	0,81208	0,79944
MAD	0,01404	0,01335	0,01417	0,01357	0,01514	0,01523	0,01418	0,01349

Tabela 4.24: Estatística dos erros da diferença da estimação sobre os conjuntos de 40 pontos.

	k_{euc}	k_{gauss}	k_{wave}	k_{wlts}	k_{pol}	k_{exp}	k_{circ}	k_{RQ}
Mínimo	0,88307	0,94415	0,96824	0,95762	0,975708	0,87479	0,88239	0,94186
Máximo	1,34502	1,54389	1,84415	1,68664	1,976296	1,26605	1,33627	1,54682
Média	1,00146	1,11322	1,19117	1,15630	1,217685	0,97771	0,99928	1,10840
D.Padrão	0,14417	0,19991	0,28298	0,23909	0,321091	0,12273	0,14173	0,19982
Mediana	0,95605	1,03601	1,08842	1,06650	1,103443	0,93958	0,95473	1,03271
MAD	0,06351	0,10253	0,13320	0,12280	0,142384	0,05318	0,06265	0,10193

Em todas elas, os números em negrito representam os núcleos que tiveram um desempenho melhor que o núcleo euclidiano, ou seja melhor que o método original. Em todas essas comparações, ficou claro novamente que a modificação proposta pode gerar resultados melhores para uma boa escolha do parâmetro θ . Na tabela 4.25 estão listados os valores de θ utilizados na comparação. Para cada núcleo k e para cada valor de $N \in \{10, 20, 30, 40\}$, o valor de θ “ótimo” foi determinado escolhendo num conjunto predeterminado de valores e aquele que produzia o menor erro médio quadrático para um dos 10 conjuntos com N amostras.

Tabela 4.25: Tabela com os valores de θ nas reconstruções com 10, 20, 30 e 40 pontos.

	10 Pontos	20 Pontos	30Pontos	40 Pontos
k_{gauss}	0,4	0,4	0,4	0,4
k_{wave}	1	1	1	1
k_{wlts}	1	1	1	1
k_{pol}	0.4	4	2	2
k_{exp}	2	2	2	2
k_{circ}	π	π	π	π
k_{RQ}	0,4	0,4	0,4	0,4

Para avaliar o tempo de execução foi escolhido comparar o tempo de avaliação do campo reconstruído no ponto $\mathbf{x} = (0.5, 0.5)$ para cada um dos

núcleos em um conjunto de entrada com $N = 30$ pontos. A tabela 4.26 mostra o resultado da comparação.

Tabela 4.26: Tempo de execução para o cálculo de $\hat{f}_k(0.5, 0.5)$ considerando um conjunto com $N = 40$ pontos.

Núcleo	Tempo	Razão
k_{euc}	0,02s	1,00s
k_{gauss}	0,03s	1,50s
k_{wave}	0,03s	1,50s
k_{wvltts}	0,03s	1,50s
k_{pol}	0,03s	1,50s
k_{exp}	0,02s	1,00s
k_{circ}	0,03s	1,50s
k_{RQ}	0,03s	1,50s

4.4

Conclusões

Nesse capítulo foi observado que a modificação proposta para o método de Shepard pode gerar resultados mais robustos do ponto de vista estatístico quando comparados com o método de Shepard original. Uma boa justificativa para isso pode ser observada ao se comparar o gráfico do quadrado da distância à origem usando o núcleo euclidiano (Figura 4.29) com o gráfico do quadrado da distância à origem com o uso do núcleo gaussiano (Figura 4.30). Observe que à medida que x se afasta da origem no núcleo euclidiano a distância cresce indefinidamente, e que no caso gaussiano ela tende a 2. Isso significa que o peso dos pontos que estão muito distantes de um dado ponto \mathbf{x}_i vai ser praticamente nulo no caso euclidiano. E no caso gaussiano, todos aqueles que estiverem fora da área de influência do núcleo que é controlado pelo valor θ vão ter um peso igual. Isso se repete em quase todos os núcleos que são estacionários e isotrópicos.

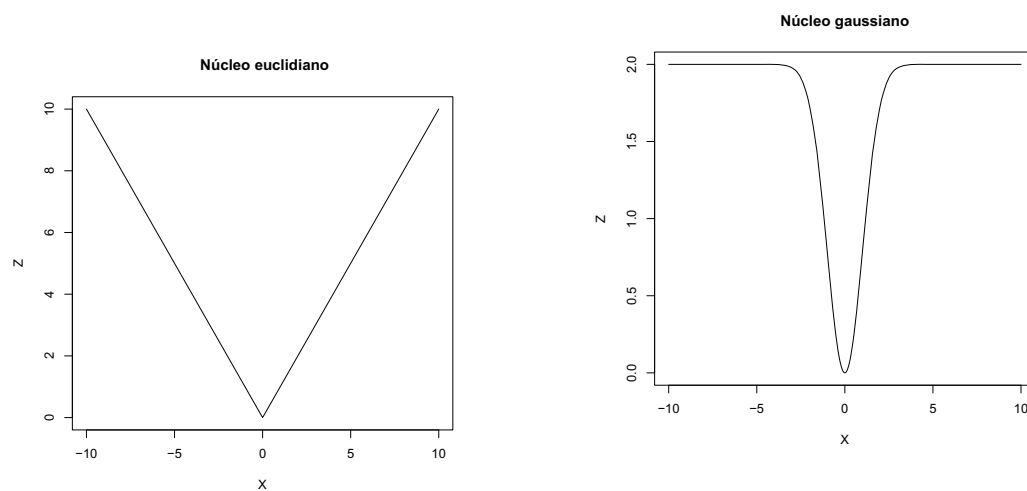


Figura 4.29: Gráfico da função distância usando o núcleo euclidiano unidimensional.

Figura 4.30: Gráfico da função distância usando o núcleo gaussiano unidimensional com parâmetro $\theta = 1$.