INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO
DEPARTAMENTO DE MATEMÁTICA APLICADA E ESTATÍSTICA

Exercícios de Mínimos Quadrados

1. **Provar que a matriz de mínimos quadrados é definida positiva, isto é, que todos os seus autovalores são estritamente positivos.**

   Seja a matriz de mínimos quadrados definida por

$$M = \begin{pmatrix} (\varphi_1, \varphi_1) & \dots & (\varphi_n, \varphi_1) \\ \vdots & & \\ (\varphi_1, \varphi_n) & \dots & (\varphi_n, \varphi_n) \end{pmatrix},$$

   em que $\varphi_1, \dots, \varphi_n$ é uma base para o espaço vetorial V e $(\cdot, \cdot)$ é o produto escalar. Seja $\lambda \in \mathbb{R}$ e $v \in V$ tal que $v \neq 0$, devemos provar que se $Mv = \lambda v$ então $\lambda > 0$.

   Premultiplicando a equação do autovetor por v^T, obtemos $v^T M v = \lambda v^T v$. Então vemos que

$$v^T A v = (v_1 \varphi_1 + \dots + v_n \varphi_n, \quad v_1 \varphi_1 + \dots + v_n \varphi_n) > 0,$$

   uma vez que $v \neq 0$ e $\{\varphi_i\}$, para $i = 1, \dots, n$, é base. Por outro lado

$$v^T v = v_1^2 + \dots v_n^2 > 0,$$

   pois $v \neq 0$, assim

$$\lambda = \frac{v^T M v}{v^T v} > 0.$$

   Portanto, todo autovalor λ da matriz M é estritamente positivo. Portanto a matriz M de mínimos quadrados é definida positiva.

2. **Calcule a distância entre duas funções f e F para cada produto escalar.**

   (a) $(u, v) = \sum_{i=1}^{M} u(x_i)\, v(x_i)$. **Resposta:**

$$\mathcal{D}(f, F) = \left[\sum_{i=1}^{N} |f(x_i) - F(x_i)|^2 \right]^{\frac{1}{2}}$$

   Portanto, minimos quadrados nesse caso minimiza a diferença quadrática das funções nos pontos de amostragem.

   (b) $(u, v) = \int_a^b u(x)\, v(x)\, dx$. **Resposta:**

$$\mathcal{D}(f, F) = \left[\int_a^b |f(x) - F(x)|^2 \, dx \right]^{\frac{1}{2}}$$

   Portanto, minimos quadrados nesse caso minimiza a integral da diferença quadrática entre ambas funções no intervalo de integração.

(c) $(u,v) = \int_a^b u(x)^2\, v(x)\, dx$. **Resposta:** Isto não é um produto escalar porque não é linear no primeiro argumento.

(d) $(u,v) = \sum_{i=1}^M W_i\, u(x_i)\, v(x_i)$. **Resposta:**

$$\mathcal{D}(f,F) = \left[\sum_{i=1}^N W_i\, |f(x_i) - F(x_i)|^2 \right]^{\frac{1}{2}}$$

Portanto, minimos quadrados nesse caso minimiza a diferença quadrática das funções nos pontos de amostragem, dando peso W_i à diferença no ponto i-éssimo.

(e) $(u,v) = \int_a^b (1+\sin^2(x))\, u(x)\, v(x)\, dx$. **Resposta:**

$$\mathcal{D}(f,F) = \left[\int_a^b (1+\sin^2(x))\, |f(x) - F(x)|^2 \; dx \right]^{\frac{1}{2}}$$

Portanto, minimos quadrados nesse caso minimiza a integral da diferença quadrática entre ambas funções no intervalo de integração, *pesada com a função* $1+\sin^2(x)$ *(e assim dando mais importância aos* x *próximos de* $x = \pi/2 + k\,\pi$*).*

3. **Seja Ω um domínio bi ou tridimensional, que está dividido em sub-domínios k_i, $i = 1, \ldots, N$ de forma arbitrária. Seja f uma função definida em Ω e seja V o espaço das funções que são constantes em cada k_i. Mostrar que a melhor aproximação de f em V, no sentido dos mínimos quadrados e considerando o produto escalar de $L^2(\Omega)$, é a função que em cada k_i vale a média de f em k_i.**

Seja f uma função definida em Ω (bidimensional por simplicidade) e seja V o espaço das funções que são constantes em cada parte k_i de Ω. Queremos mostrar que a melhor aproximação de f em V, no sentido dos mínimos quadrados e considerando o produto escalar $L^2(\Omega)$, é a função que em cada k_i vale a média de f em k_i. Lembremos ante tudo que o produto escalar de $L^2(\Omega)$ entre duas funções u e v é

$$(u,v) \;=\; \int_\Omega u(x,y)\, v(x,y)\; dx\, dy$$

Seja φ_i a função que vale 1 em k_i e zero fora dele, isto é

$$\varphi_i(x,y) = \begin{cases} 1 & \text{se}\,(x,y) \in\, k_i \\ 0 & \text{se não} \end{cases}$$

Assim aproximamos a função f por

$$f \approx F^* = \alpha_1^* \varphi_1 + \ldots + \alpha_N^* \varphi_N \tag{1}$$

no sentido dos mínimos quadrados, onde N é o número de partes (ou subdomínios). Queremos resolver o sistema $M\alpha^* = B$ em que

$$M = \begin{bmatrix} (\varphi_1,\varphi_1) & \dots & (\varphi_n,\varphi_1) \\ \vdots & & \\ (\varphi_1,\varphi_n) & \dots & (\varphi_n,\varphi_n) \end{bmatrix}, \quad \alpha^* = \begin{bmatrix} \alpha_1^* \\ \vdots \\ \alpha_n^* \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} (\varphi_1, f) \\ \vdots \\ (\varphi_n, f) \end{bmatrix}$$

Aplicando o produto escalar definido acima, temos

$$\begin{aligned} (\varphi_i,\varphi_i) &= \textstyle\int_\Omega \varphi_i(x,y)\,\varphi_i(x,y)\,dx\,dy \\ &= \underbrace{\int_{k_1} \varphi_i\varphi_i dxdy}_{0} + \ldots + \underbrace{\int_{k_i} \varphi_i\varphi_i dxdy}_{\int_{k_i} 1^2 dxdy} + \ldots + \underbrace{\int_{k_n} \varphi_i\varphi_i dxdy}_{0} \\ &= \textstyle\int_{k_i} 1^2\,dx\,dy \;=\; \text{Area}(k_i) \end{aligned}$$

e, se $j \neq i$,

$$\begin{aligned} (\varphi_i,\varphi_j) &= \textstyle\int_\Omega \varphi_i\varphi_j dxdy \\ &= \underbrace{\int_{k_1} \varphi_i\varphi_j dxdy}_{0} + \ldots + \underbrace{\int_{k_i} \varphi_i\varphi_j dxdy}_{0} + \ldots + \underbrace{\int_{k_j} \varphi_i\varphi_j dxdy}_{0} + \ldots + \underbrace{\int_{k_n} \varphi_i\varphi_i dxdy}_{0} \\ &= 0 \end{aligned}$$

Ademais,

$$\begin{aligned} (\varphi_i, f) &= \textstyle\int_\Omega \varphi_i(x,y)\,f(x,y)dxdy \\ &= \underbrace{\int_{k_1} \varphi_i\, f\,dx\,dy}_{0} + \ldots + \underbrace{\int_{k_i} \varphi_i f\,dx\,dy}_{\int_{k_i} f(x,y)dxdy} + \ldots + \underbrace{\int_{k_n} \varphi_i f\,dx\,dy}_{0} \\ &= \textstyle\int_{k_i} f(x,y)\,dx\,dy \end{aligned}$$

assim a matriz M e o lado direito B do sistema são dados por

$$M = \begin{bmatrix} \text{Area}(k_1) & 0 & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ 0 & \dots & \text{Area}(k_i) & 0 \\ 0 & \dots & \ddots & 0 \\ 0 & \dots & 0 & \text{Area}(k_N) \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} \int_{k_1} f(x,y)\;dx\,dy \\ \vdots \\ \int_{k_i} f(x,y)\;dx\,dy \\ \vdots \\ \int_{k_N} f(x,y)\;dx\,dy \end{bmatrix}$$

Resolvemos o sistema facilmente por ser M diagonal, determinando os coeficientes α_i^* que são dados por

$$\alpha_i^* = \frac{1}{\text{Area}(k_i)} \int_{k_i} f(x,y)\,dx\,dy$$

Assim a função F^* em cada k_i vale a média de f em k_i. Para um domínio tridimensional a resolução se faz de forma análoga. Refaça o exercício considerando agora que $\varphi_i(x, y)$ vale i no subdomínio k_i e zero fora dele. Comprove que, embora M, B e α^* resultam diferentes, a função F^* obtida é exatamente a mesma.

4. **Considere os seguintes conjuntos de funções:**
   V: os polinômios de grau ≤ 4 tais que $p(1) = 0$;
   W: o subconjunto de V tal que $p'(2) = 0$;
   X: o subconjunto de W tal que $p''(3) = 1$.
   São V, W, X espaços vetoriais? De que dimensão?

   Resposta: V é espaço vetorial de dimensão 4.
   W é espaço vetorial de dimensão 3.
   X não é espaço vetorial, porque se $p \in X$ e $q \in X$, então $(p + q)''(3) = 2$ e portanto $p + q$ não pertence a X. Outro bom motivo é que o polinômio nulo não pertence a X.

5. **Sete experimentos foram feitos para determinar a durabilidade de um certo tipo de pneus. Os resultados foram os seguintes:**

Experimento (i)	1	2	3	4	5	6	7
Durabilidade (D_i) $[km]$	56000	52000	55000	62000	60000	63000	61000

Pelas condições dos experimentos, os engenheiros decidiram associar a cada um deles um peso (fiabilidade) diferente:

$$w_1 = 1, \quad w_2 = 1.5, \quad w_3 = 2, \quad w_4 = 2.5, \quad w_5 = 3, \quad w_6 = 4, \quad w_7 = 5.$$

Calcular a melhor estimação D^*, no sentido de mínimos quadrados, para a duravilidade dos pneus considerados utilizando o produto escalar

$$(\underline{D}, \underline{G}) = \sum_{i=1}^{7} w_i D_i G_i. \tag{2}$$

Como o que se procura é um número único independente dos experimentos (a "durabilidade nominal" do pneu ou algum nome assim), a função que procuramos é uma constante da forma

$$F = \alpha_1$$

e por tanto $\varphi_1 = 1$. Assim a melhor estimação F^*, no sentido de mínimos quadrados, para durabilidade dos pneus é dada por

$$F^* = \alpha_1^* \tag{3}$$

Considerando o produto escalar (2) o sistema a resolver (que é de 1 equação com 1 incógnita) é

$$M = [(1, 1)] = \sum_{i=1}^{7} w_i = 19; \qquad B = [(1, f)] = \sum_{i=1}^{7} w_i \, D_i$$

substituindo os valores obtemos que a durabilidade procurada é

$$F^* = 59789,4736$$

6. **Seja a função $f(x)$ tal que $f(x) = 1$ se $x < 3$ e $f(x) = 3$ se $x > 3$. Calcular a melhor aproximação $F(x)$ dessa função, no sentido dos mínimos quadrados, no intervalo (0.4), dentre os polinômios de primeiro grau. O produto escalar é $(f, g) = \int_0^4 f(x)\, g(x)\, dx$.**

   Aproximamos $f(x)$ por

$$f(x) \approx F^*(x) = \alpha_1^* + \alpha_2^* x \tag{4}$$

   e portanto $\varphi_1(x) = 1$ e $\varphi_2(x) = x$.

   Agora calculamos

$$\begin{aligned}
(\varphi_1, \varphi_1) &= \int_0^4 1\ dx \ = 4 \\
(\varphi_1, \varphi_2) &= \int_0^4 x\ dx \ = 8 \\
(\varphi_2, \varphi_2) &= \int_0^4 x^2\ dx \ = \frac{64}{3} \\
(\varphi_1, f) &= \int_0^4 f(x)\ dx = \int_0^3 1\ dx + \int_3^4 3\ dx \ = 6 \\
(\varphi_2, f) &= \int_0^4 x\, f(x)\ dx = \int_0^3 x\ dx + \int_3^4 3x\ dx \ = 15
\end{aligned}$$

   Assim chegamos ao sistema

$$\begin{pmatrix} 4 & 8 \\ 8 & 64/3 \end{pmatrix} \begin{pmatrix} \alpha_1^* \\ \alpha_2^* \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \end{pmatrix}$$

   cujo resultado é $\alpha_1^* = 0.375$ e $\alpha_2^* = 0.5625$. Assim,

$$F^*(x) = 0.375 + 0.5625\, x$$

7. **Considerando a função $y = f(x)$ dada pela tabela:**

x_i	-1	0	1	3
y_i	2.1	2.9	3	5.7

   Ajustá-la por um polinômio de primeiro grau, usando o método dos mínimos quadrados.

   Uma base para o conjunto dos polinômios de primeiro grau é dada por $\{1, x\}$, assim vamos aproximar $f(x)$ por

$$f(x) \approx F^*(x) = \alpha_1^* + \alpha_2^* x$$

considerando o produto escalar

$$(u, v) = \sum_{i=1}^{4} u(x_i)\, v(x_i)$$

Para isto vemos que

$$\begin{aligned}
(1,1) &= 4 \\
(1,x) &= 1 \times (-1) + 1 \times 0 + 1 \times 1 + 1 \times 3 = 3 \\
(x,x) &= (-1) \times (-1) + 0 \times 0 + 1 \times 1 + 3 \times 3 = 11 \\
(1,f) &= 1 \times 2.1 + 1 \times 2.9 + 1 \times 3 + 1 \times 5.7 = 13.7 \\
(x,f) &= (-1) \times 2.1 + 0 \times 2.9 + 1 \times 3 + 3 \times 5.7 = 18
\end{aligned}$$

e portanto o sistema a resolver é

$$\begin{bmatrix} 4 & 3 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \end{bmatrix} = \begin{bmatrix} 13.7 \\ 18 \end{bmatrix} \tag{5}$$

A solução do sistema linear é dada por

$$\alpha_1^* = 2.76286 \quad \alpha_2^* = 0.88286$$

e assim

$$f(x) \approx F^*(x) = 2.76286 + 0.88286\, x.$$

8. **Considere a função**

$$f(x) = cos(x)$$

Encontre a melhor aproximação polinomial $p(x)$ **de segundo grau, no sentido de mínimos quadrados, para a função f no intervalo** $[0, \frac{\pi}{2}]$**. Utilize o produto escalar** $(f, g) = \int_0^{\frac{\pi}{2}} f(x)g(x)dx$**.**

Queremos aproximar $f(x)$ da seguinte forma

$$f(x) \approx F^*(x) = p(x) = \alpha_1^* + \alpha_2^* x + \alpha_3^* x^2$$

no sentido dos mínimos quadrados, com base definida por $\{\varphi_1, \varphi_2, \varphi_3\} = \{1, \quad x, \quad x^2\}$. Para isso resolvemos o seguinte sistema

$$\begin{bmatrix} (\varphi_1, \varphi_1) & (\varphi_2, \varphi_1) & (\varphi_3, \varphi_1) \\ (\varphi_1, \varphi_2) & (\varphi_2, \varphi_2) & (\varphi_3, \varphi_2) \\ (\varphi_1, \varphi_3) & (\varphi_2, \varphi_3) & (\varphi_3, \varphi_3) \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \\ \alpha_3^* \end{bmatrix} = \begin{bmatrix} (f, \varphi_1) \\ (f, \varphi_2) \\ (f, \varphi_3) \end{bmatrix}$$

com

$$
\begin{aligned}
(\varphi_1, \varphi_1) &= \int_0^{\frac{\pi}{2}} 1 \quad 1 \quad dx = \frac{\pi}{2} \\
(\varphi_1, \varphi_2) &= \int_0^{\frac{\pi}{2}} 1 \quad x \quad dx = \frac{\pi^2}{8} \\
(\varphi_1, \varphi_3) &= \int_0^{\frac{\pi}{2}} 1 \quad x^2 \quad dx = \frac{\pi^3}{24} \\
(\varphi_2, \varphi_3) &= \int_0^{\frac{\pi}{2}} x \quad x^2 \quad dx = \frac{\pi^4}{64} \\
(\varphi_2, \varphi_2) &= \int_0^{\frac{\pi}{2}} x \quad x \quad dx = \frac{\pi^3}{24} \\
(\varphi_3, \varphi_3) &= \int_0^{\frac{\pi}{2}} x^2 \quad x^2 \quad dx = \frac{\pi^5}{160} \\
(f, \varphi_1) &= \int_0^{\frac{\pi}{2}} \cos x dx = 1 \\
(f, \varphi_2) &= \int_0^{\frac{\pi}{2}} x \quad \cos x dx = \frac{\pi - 2}{2} \\
(f, \varphi_3) &= \int_0^{\frac{\pi}{2}} \cos x \quad x^2 \quad dx = \frac{\pi^2 - 8}{4}
\end{aligned}
$$

temos então,

$$
\begin{bmatrix} \frac{\pi}{2} & \frac{\pi^2}{8} & \frac{\pi^3}{24} \\ \frac{\pi^2}{8} & \frac{\pi^3}{24} & \frac{\pi^4}{64} \\ \frac{\pi^3}{24} & \frac{\pi^4}{64} & \frac{\pi^5}{160} \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \\ \alpha_3^* \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\pi-2}{2} \\ \frac{\pi^2-8}{4} \end{bmatrix}
$$

Resolvendo o sistema acima, com $\pi = 3.14$, obtemos $\alpha_1^* = 0.9722$, $\alpha_2^* = 0.0019$ e $\alpha_3^* = -0.4152$. Assim $f(x)$ é aproximada por

$$f(x) \approx p(x) = 0.9722 + 0.0019x + 0.4152x^2.$$

9. **Seja o seguinte sistema de equações:**

$$
\begin{aligned}
x + 2y + z &= -5.4 \\
2x - 2y &= 9 \\
x + z &= 2.2 \\
4x + 6z &= 8.5 \\
x - z &= 1.8 \\
6x + y - z &= 4.7
\end{aligned}
\tag{6}
$$

   (a) **Calcular a "melhor solução" no sentido dos mínimos quadrados.**

Escrevendo o sistema linear incompatível dado como $A\alpha = b$, queremos resolver o sistema $M\alpha^* = B$, com $M = A^TCA$ e $B = A^TCb$. Temos que

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -2 & 0 \\ 1 & 0 & 1 \\ 4 & 0 & 6 \\ 1 & 0 & -1 \\ 6 & 1 & -1 \end{bmatrix}$$

e

$$b = \begin{bmatrix} -5.4 \\ 9 \\ 2.2 \\ 8.5 \\ 1.8 \\ 4.7 \end{bmatrix}$$

Tomando C como a matriz identidade (porque sim), obtem-se $x = \alpha_1^* = 1.4211$, $y = \alpha_2^* = -3.3584$ e $z = \alpha_3^* = 0.4414$.

(b) **Qual seria a "melhor solução" para x e y se você tivesse certeza absoluta de que $z = 0.5$ e como a calcularia?**

Consideramos $z = 0.5$ e reescrevemos o sistema (6) da seguinte forma:

$$\begin{aligned} x + 2y &= -5.9 \\ 2x - 2y &= 9 \\ x &= 1.7 \\ 4x &= 5.6 \\ x &= 2.3 \\ 6x + y &= 5.2 \end{aligned}$$

Resolvemos o sistema $M\alpha^* = B$, com $M = A^TCA$ e $B = A^TCb$, em que agora

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \\ 1 & 0 \\ 4 & 0 \\ 1 & 0 \\ 6 & 1 \end{bmatrix}$$

e

$$b = \begin{bmatrix} -5.9 \\ 9 \\ 1.7 \\ 5.6 \\ 2.3 \\ 5.2 \end{bmatrix}$$

novamente escolhendo C como a matriz identidade (porque não, né?) obtemos $x = \alpha_1^* = 1.4091$ e $y = \alpha_2^* = -3.3596$.

10. **Seja o seguinte sistema de equações:**

$$\begin{aligned} x + y + z &= -1.4 \\ 3x - y &= 7.5 \\ x + z &= 2.2 \\ 4x + 6z &= 6.5 \\ 7x - 5z &= 7.8 \\ x + y - z &= 4.7 \end{aligned}$$

Calcular a solução no sentido dos mínimos quadrados.

Resolvendo como no exercício anterior achamos $x = \alpha_1^* = 2.7315$, $y = \alpha_2^* = 0.9708$ e $z = \alpha_3^* = -2.1904$.

11. **Calcular, tanto por mínimos quadrados quanto por interpolação o polinômio quadrático que ajusta os três pares (x, y) seguintes: $(0, 4)$, $(3, 2)$, $(4, 4)$. Discuta as diferenças e semelhanças entre os polinômios obtidos pelos dois métodos.**

Vamos considerar os pontos

x	0	3	4
y	4	2	4

Minimos Quadrados

Vamos aproximar os dados da tabela da seguinte forma

$$p(x) = \alpha_1^* + \alpha_2^* x + \alpha_3^* x^2.$$

no sentido dos mínimos quadrados, considerando o produto

$$(u, v) = u(0)v(0) + u(3)v(3) + u(4)v(4)$$

Resolvemos o seguinte sistema

$$\begin{bmatrix} 3 & 7 & 25 \\ 7 & 25 & 91 \\ 25 & 91 & 337 \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \\ \alpha_3^* \end{bmatrix} = \begin{bmatrix} 10 \\ 22 \\ 82 \end{bmatrix}$$

obtendo assim $\alpha_1^* = 4$, $\alpha_2^* = -2.6$ e $\alpha_3^* = 0.6$ e portanto

$$p(x) = 4 - 2.6x + 0.6x^2.$$

Interpolação

Vamos aproximar os dados da tabela por interpolação, ou seja,

$$p(x) = a + bx + cx^2.$$

Resolvemos o sistema (matriz de van der Monde)

$$\begin{cases} a + 0b + 0c &= 4 \\ a + 3b + 9c &= 2 \\ a + 4b + 16c &= 4 \end{cases}$$

e obtemos $a = 4$, $b = -2.6$ e $c = 0.6$. e portanto

$$p(x) = 4 - 2.6x + 0.6x^2.$$

Observações: O polinômio encontrado tanto por métodos dos mínimos quadrado quanto por interpolação quadrática é o mesmo. Isto acontece porque o polinômio interpolador pertence ao espaço V em que são realizados os mínimos quadrados (o espaço de polinômios quadráticos). Sendo que o polinômio interpolador tem erro zero (distância zero), ele necessariamente coincide com o resultado dos mínimos quadrados. Notar que se se acrescentasse mais um ponto, em geral não haveria um polinômio quadrático que passasse exatamente pelos quatro pontos. Nesse caso só restaria mínimos quadrados como procedimento de aproximação com quadráticas. Se, mesmo com quatro pontos, desse a casualidade que eles podem ser interpolados com um polinômio quadrático (lembremos que "interpolar" implica passar exatamente pelos pontos), então novamente os resultados da interpolação e dos mínimos quadrados seriam coincidentes. **Conclusão:** Alguém que realmente sabe a matéria não teria calculado por ambos métodos, simplesmente teria calculado por um deles e afirmado com total certeza que o outro ia dar o mesmo resultado.