Soluções para problemas selecionados da apostila
Otimização Matemática e Pesquisa Operacional
de André R. Fioravanti e Matheus Souza

– Capítulo 3 –

Plínio Santini Dester (p103806@dac.unicamp.br)

18 de janeiro de 2022

*Em caso de dúvidas, sugestões ou correções, não hesite em mandar um e-mail.

3 Problemas

3.1. Determine a solução de quadrados mínimos do sistema sobredeterminado

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 3 \\ x_1 + 3x_2 = 6 \\ x_1 + 4x_2 = 10 \end{cases}$$

Represente graficamente as retas que definem este sistema e a solução obtida. Interprete graficamente.

Solução: Queremos resolver o problema $\min_{x\in\mathbb{R}^2} ||Ax - b||_2^2$, onde

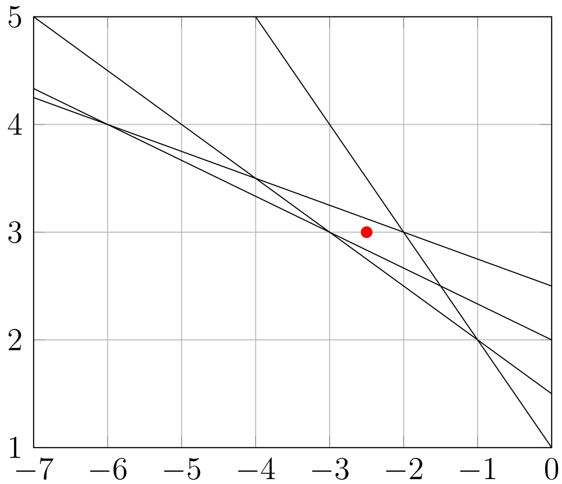
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 10 \end{bmatrix}.$$

A solução é dada pela equação normal, ou seja, a solução x^* deve satisfazer

$$A^\top A x^* = A^\top b.$$

Dessa forma,

$$\begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} x^* = \begin{bmatrix} 20 \\ 65 \end{bmatrix} \Rightarrow x^* = \frac{1}{120 - 100} \begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix} \begin{bmatrix} 20 \\ 65 \end{bmatrix}$$
$$\Rightarrow \boxed{x^* = \begin{bmatrix} -5/2 \\ 3 \end{bmatrix}}.$$



Podemos ver pelo gráfico que a solução aparenta ser o ponto no qual as retas se encontram próximas entre si e ainda está próximo das 6 soluções entre pares de equações.

3.2. Considere o problema de quadrados mínimos $\min \|Ax - b\|_2^2$ com

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Determine a solução ótima $x^\star$ deste problema. Verifique que $r^\star = b - Ax^\star$ é ortogonal às colunas de A. E se o vetor b for trocado por $\hat{b} = [1\ 3\ 4]^\top$, o que acontece com a solução ótima e o resíduo associado? Interprete.

Solução: Novamente, a solução é dada pela equação normal, ou seja, a solução x^* deve satisfazer
$$A^\top A x^* = A^\top b.$$

Dessa forma,

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow x^* = \frac{1}{4-1} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\Rightarrow \boxed{x^* = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}}.$$

Assim, $r^* = [2/3, 2/3, -2/3]^\top$. Como $A^\top r^* = [0, 0]^\top$, então r^* é, de fato, ortogonal às colunas de A. Se b for trocado por $\hat{b}$, então o sistema de equações passa a ter solução, ou seja, o resíduo associado será nulo.

3.3. Encontre a projeção de b em $\mathcal{R}(A)$, sendo

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}.$$

Decomponha b como $p + q$, com $p \in \mathcal{R}(A)$ e q ortogonal à $\mathcal{R}(A)$. Em que subespaço está q?

Solução: A projeção de b no conjunto $\mathcal{R}(A)$ pode ser encontrada através do seguinte problema de otimização $\min_{x \in \mathbb{R}^2} ||Ax - b||_2^2$, de forma que a projeção é dada por $p = Ax^*$. Como visto anteriormente, a solução deve satisfazer a equação normal, ou seja, $A^\top A\, p = A^\top b$. Assim,

$$\begin{bmatrix} 9 & -9 \\ -9 & 18 \end{bmatrix} x^* = \begin{bmatrix} -9 \\ 27 \end{bmatrix} \Rightarrow x^* = \frac{1}{162-81} \begin{bmatrix} 18 & 9 \\ 9 & 9 \end{bmatrix} \begin{bmatrix} -9 \\ 27 \end{bmatrix}$$
$$\Rightarrow \boxed{x^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}}.$$

Portanto, $p = Ax^* = [3, 0, 6]^\top$ e, por definição, $q = b - p = [-2, 2, 1]^\top$. Como q é ortogonal às colunas de A, pois $A^\top q = [0, 0]^\top$, então $q \in \mathcal{N}(A^\top) = \mathcal{R}(A)^\perp$.

3.4. **Fórmulas de Regressão Linear** Considere o problema em que um conjunto de dados $(t^{(i)}, y^{(i)})$, $i = 1, \cdots, N$, deve ser ajustado, no sentido de quadrados mínimos, por uma reta

$$\phi(t) = x_1 + x_2 t.$$

Construa o sistema normal associado a este problema e resolva-o para mostrar que os coeficientes ótimos são dados por

$$x_1^\star = \frac{\left(\sum_{i=1}^N \left(t^{(i)}\right)^2\right)\left(\sum_{i=1}^N y^{(i)}\right) - \left(\sum_{i=1}^N t^{(i)}\right)\left(\sum_{i=1}^N t^{(i)} y^{(i)}\right)}{N\sum_{i=1}^N \left(t^{(i)}\right)^2 - \left(\sum_{i=1}^N t^{(i)}\right)^2}$$

e

$$x_2^\star = \frac{N\left(\sum_{i=1}^N t^{(i)} y^{(i)}\right) - \left(\sum_{i=1}^N t^{(i)}\right)\left(\sum_{i=1}^N y^{(i)}\right)}{N\sum_{i=1}^N \left(t^{(i)}\right)^2 - \left(\sum_{i=1}^N t^{(i)}\right)^2}.$$

Solução: Para obter o melhor ajuste dos coeficientes x_1, x_2 devemos resolver o seguinte problema de otimização

$$\min_{x\in\mathbb{R}^2} \sum_{i=1}^N ||\phi(t^{(i)}) - y^{(i)}||_2^2 \;\sim\; \min_{x\in\mathbb{R}^2} ||\Phi x - y||_2^2,$$

onde

$$\Phi = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t^{(1)} & t^{(2)} & \cdots & t^{(N)} \end{bmatrix}^\top .$$

Logo, podemos usar a equação normal

$$\begin{aligned}
x^* &= (\Phi^\top \Phi)^{-1}\Phi^\top y \\
&= \begin{bmatrix} N & \sum_{i=1}^N t^{(i)} \\ \sum_{i=1}^N t^{(i)} & \sum_{i=1}^N (t^{(i)})^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^N y^{(i)} \\ \sum_{i=1}^N t^{(i)} y^{(i)} \end{bmatrix} \\
&= \frac{1}{N\sum_{i=1}^N (t^{(i)})^2 - \left(\sum_{i=1}^N t^{(i)}\right)} \begin{bmatrix} \sum_{i=1}^N (t^{(i)})^2 & -\sum_{i=1}^N t^{(i)} \\ -\sum_{i=1}^N t^{(i)} & N \end{bmatrix} \begin{bmatrix} \sum_{i=1}^N y^{(i)} \\ \sum_{i=1}^N t^{(i)} y^{(i)} \end{bmatrix} \\
&= \frac{1}{N\sum_{i=1}^N (t^{(i)})^2 - \left(\sum_{i=1}^N t^{(i)}\right)} \begin{bmatrix} \left(\sum_{i=1}^N \left(t^{(i)}\right)^2\right)\left(\sum_{i=1}^N y^{(i)}\right) - \left(\sum_{i=1}^N t^{(i)}\right)\left(\sum_{i=1}^N t^{(i)} y^{(i)}\right) \\ N\left(\sum_{i=1}^N t^{(i)} y^{(i)}\right) - \left(\sum_{i=1}^N t^{(i)}\right)\left(\sum_{i=1}^N y^{(i)}\right) \end{bmatrix}.
\end{aligned}$$

3.10. **Solução de Quadrados Mínimos de Norma Mínima.** Quando a matriz $A \in \mathbb{R}^{m\times n}$ do problema de quadrados mínimos

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2,$$

com $b \in \mathbb{R}^m$, não tiver posto completo, este problema não tem solução única. Neste caso, pode ser interessante encontrar o vetor $x^\star$ que é solução ótima deste problema e que tem norma Euclidiana mínima. Usando-se o sistema normal, podemos escrever este problema como

$$\begin{array}{rl} \min & \|x\|_2^2 \\ \text{sujeito a} & A^\top A x = A^\top b. \end{array}$$

A restrição de igualdade deste problema pode ser removida parametrizando-se o núcleo de $A^\top A$; assim, obtemos um problema de otimização irrestrita.

Encontre a solução de norma Euclidiana mínima para os problemas de quadrados mínimos definidos pelas matrizes abaixo:

(a) $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$;

(b) $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ -1 & -2 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Solução: Para ambos itens, podemos resolver da seguinte maneira. Seja F a matriz formada pelos vetores que geram $\mathcal{N}(A^\top A)$, então um problema de otimização equivalente pode ser enunciado como

$$\min_{z \in \mathbb{R}^k} ||\bar{x} + Fz||_2^2 \sim \min_{z \in \mathbb{R}^k} ||Fz - (-\bar{x})||_2^2,$$

onde k é n menos o posto da matriz $A^T A$ e $\bar{x}$ é uma solução particular do sistema de equações normais. Note que podemos colocar esse problema equivalente como um problema de quadrados mínimos e resolvê-lo usando a equação normal, ou seja,

$$F^\top F z^* = F^\top(-\bar{x}).$$

(a) Montando o sistema de equações normais $A^\top A x = A^\top b$, temos que

$$\begin{bmatrix} 6 & 2 & 4 \\ 2 & 2 & 0 \\ 4 & 0 & 4 \end{bmatrix} x = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

Nesse problema, $k = 1$, logo, seja $x_1 = z$, então da segunda linha de equações, temos que $x_2 = -z$ e da terceira linha $x_3 = 1/2 - z$. Logo,

$$x = \begin{bmatrix} z \\ -z \\ 1/2 - z \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1/2 \end{bmatrix}}_{\bar{x}} + \underbrace{\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}}_{F} z.$$

Assim, $z^* = (F^\top F)^{-1} F^\top(-\bar{x}) = (1/3)(1/2) = 1/6$.

Portanto, $\boxed{x^* = [-1/6, 1/6, 1/3]^\top}$.

(b) Montando o sistema de equações normais $A^\top A x = A^\top b$, temos que

$$\begin{bmatrix} 6 & 12 \\ 12 & 24 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Nesse problema, $k = 1$, logo, seja $x_2 = z$, então da primeira linha de equações, temos que $x_1 = 1/6 - 2z$. Logo,

$$x = \begin{bmatrix} 1/6 - 2z \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 1/6 \\ 0 \end{bmatrix}}_{\bar{x}} + \underbrace{\begin{bmatrix} -2 \\ 1 \end{bmatrix}}_{F} z.$$

Assim, $z^* = (F^\top F)^{-1} F^\top (-\bar{x}) = (1/5)(1/3) = 1/15$.
Portanto, $\boxed{x^* = [1/30, 1/15]^\top}$.

3.11. **Sistemas Subdeterminados: Solução de Norma Mínima e Problemas de Quadrados Mínimos com Regularização de Tikhonov.** O problema discutido acima está relacionado com o seguinte problema: dado um sistema linear $Ax = b$, com $A \in \mathbb{R}^{m \times n}$ e $b \in \mathbb{R}^m$ e $m < n$. Se este sistema for compatível, então ele admite infinitas soluções; vamos assumir que o posto de A é completo, o que assegura essa propriedade. Dentre todas estas soluções, desejamos determinar a que possui a menor norma, ou seja, desejamos resolver o problema

$$\begin{array}{rl} \min & \|x\|_2^2 \\ \text{sujeito a} & Ax = b. \end{array}$$

(a) Interprete este problema geometricamente.

(b) Mostre que a solução ótima deste problema é dada por $x^\star = A^\top (AA^\top)^{-1} b$; para tanto, tome outra solução arbitrária deste sistema $x = x^\star + z$, com z tal que $Az = 0$ e mostre que a norma é maior que ou igual à norma de $x^\star$.

(c) Este problema tem uma relação interessante com o problema de quadrados mínimos com *regularização de Tikhonov*:

$$\min \|Ax - b\|_2^2 + \lambda \|x\|_2^2,$$

em que $\lambda > 0$ é um parâmetro definido pelo projetista. A regularização de Tikhonov evita que os coeficientes de ajuste do problema de quadrados mínimos sejam arbitrariamente grandes, impondo um certo *trade-off* entre a norma do resíduo e a norma da solução.

Use as condições de otimalidade para determinar a solução ótima deste problema. O que acontece com esta solução se $\lambda \to 0$?

Solução:

(a) Esse problema é equivalente à encontrar o ponto do conjunto $\{x \in \mathbb{R}^n \mid Ax = b\}$ que esteja mais próximo à origem, ou seja, é a projeção da origem nesse conjunto.

(b) *Demonstração.* Primeiramente mostramos que, de fato, $Ax^* = b$.

$$Ax^* = AA^\top(AA^\top)^{-1}b = b,$$

pois $AA^\top$ é inversível, dado que A tem posto completo. Agora mostremos que $||x^*||_2 \leq ||x||_2$ para todo x que satisfaaça $Ax = b$. Seja $z \in \mathcal{N}(A)$,

$$\begin{aligned}
||x||_2^2 &= ||x^* + z||_2^2 \\
&= (x^* + z)^\top (x^* + z) \\
&= (x^*)^\top x^* + 2(x^*)^\top z + z^\top z \\
&= ||x^*||_2^2 + 2(A^\top(AA^\top)^{-1}b)^\top z + ||z||_2^2 \\
&= ||x^*||_2^2 + 2((AA^\top)^{-1}b)^\top \underbrace{Az}_{0} + ||z||_2^2 \\
&= ||x^*||_2^2 + ||z||_2^2 \\
&\geq ||x^*||_2^2.
\end{aligned}$$

■

(c) A função objetivo pode ser re-escrita como

$$\begin{aligned}
f_0(x) &= (Ax - b)^\top (Ax - b) + \lambda x^\top x \\
&= x^\top A^\top A x - 2x^\top A^\top b + b^\top b + \lambda x^\top x \\
&= x^\top (A^\top A + \lambda I_n) x - 2x^\top A^\top b + b^\top b,
\end{aligned}$$

onde I_n é a matriz identidade. Portanto, o gradiente é dado por

$$\nabla f_0(x) = 2(A^\top A + \lambda I_n)x - 2A^\top b.$$

Resolvendo $\nabla f_0(x^*) = 0$, obtemos

$$x^* = (A^\top A + \lambda I_n)^{-1} A^\top b.$$

Note que a matriz $(A^\top A + \lambda I_n)$ é inversível e definida positiva, dado que $\det(A^\top A) = 0$ e $\lambda > 0$. Esse problema é convexo, então não é necessário checar as condições de segunda ordem.

Como esperado, se $\lambda \to 0$, x^* converge para a solução do item anterior. O motivo é que estamos dando um peso grande para o termo $||Ax - b||_2^2$ ao dar um peso pequeno para $||x||_2^2$, ou seja, o problema se torna equivalente à minimizar $||x||_2^2$ restrito à $Ax = b$. É importante notar que λ não pode ser igual à zero, pois ao contrário de $AA^\top$, a matriz $A^\top A$ **não** é inversível.

3.14. Considere o problema de quadrados mínimos $\|b - Ax\|_2^2$ com A e b dados por

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1+\alpha \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

(a) Para que valores de $\alpha \in \mathbb{R}$ este problema possui solução única?

(b) Assumindo que $\alpha \neq 0$, resolva o problema de quadrados mínimos dado acima e calcule o resíduo ótimo. O resíduo depende de α?

(c) Investigue o que acontece com a solução $x^\star$ encontrada acima quando $\alpha \to 0$. O que acontece com o resíduo ótimo?

Solução:

(a) Usando a equação normal $A^\top A x^* = A^T b$, temos que

$$\begin{bmatrix} 2 & 2+\alpha \\ 2+\alpha & 1+(1+\alpha)^2 \end{bmatrix} x^* = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

Esse problema tem solução única se, e somente se, $\det(A^\top A) \neq 0$. Como $\det(A^\top A) = \alpha^2$, então o problema tem solução única para todo $\alpha \neq 0$.

(b) Resolvendo a equação normal, temos que

$$\begin{aligned} x^* &= (A^\top A)^{-1} A^\top b \\ &= \frac{1}{\alpha^2} \begin{bmatrix} 2+2\alpha+\alpha^2 & -(2+\alpha) \\ -(2+\alpha) & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} \\ &= \frac{2}{\alpha} \begin{bmatrix} 1+\alpha \\ -1 \end{bmatrix}. \end{aligned}$$

O resíduo ótimo é dado por

$$r^* = b - Ax^* = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{\alpha} \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Assim, o resíduo $||r^*(\alpha)||_2 = 1$ coincidentemente não depende de α.

(c) Esse limite não existe, pois $\lim_{\alpha\to 0} 1/\alpha$ é indeterminado. Quando $\alpha = 0$ a solução tem 1 grau de liberdade, já que a solução da equação normal não é única para esse caso, então era de se esperar que o limite não deveria existir.

Em relação ao resíduo, o limite existe $\lim_{\alpha\to 0} ||r^*(\alpha)||_2 = 1$. Porém, ele é diferente do resíduo quando resolvemos o problema para $\alpha = 0$, em que encontraríamos $||r^*(0)||_2 = 2$. Dessa forma, $\lim_{\alpha\to 0} ||r^*(\alpha)||_2 \neq ||r^*(0)||_2$.