Método dos Mínimos Quadrados (MMQ)

MMQ e Famílias Ortogonais

Nelson Kuhl

IME/USP

24 de setembro de 2020

MMQ Linear - Formulação Geral

Para a formulação geral do MMQ linear temos:

MMQ Linear - Formulação Geral

Para a formulação geral do MMQ linear temos:

- um espaço vetorial real V munido de um produto interno $\langle\ ,\ \rangle$, possivelmente degenerado ($\langle u, u\rangle \geq 0$, $\forall u \in V$, mas pode se anular para aluguns vetores não nulos);

MMQ Linear - Formulação Geral

Para a formulação geral do MMQ linear temos:

- um espaço vetorial real V munido de um produto interno $\langle \ , \ \rangle$, possivelmente degenerado ($\langle u, u \rangle \geq 0$, $\forall u \in V$, mas pode se anular para aluguns vetores não nulos);
- um subespaço vetorial $G \subset V$ de dimensão finita $m + 1$ tal que a restrição de $\langle \ , \ \rangle$ a ele é um produto interno ($\langle u, u \rangle > 0$, $\forall u \neq 0$ em G).

O problema é então formulado como:

MMQ Linear - Formulação Geral

Para a formulação geral do MMQ linear temos:

- um espaço vetorial real V munido de um produto interno $\langle\ ,\ \rangle$, possivelmente degenerado ($\langle u, u\rangle \geq 0$, $\forall u \in V$, mas pode se anular para aluguns vetores não nulos);
- um subespaço vetorial $G \subset V$ de dimensão finita $m+1$ tal que a restrição de $\langle\ ,\ \rangle$ a ele é um produto interno ($\langle u, u\rangle > 0$, $\forall u \neq 0$ em G).

O problema é então formulado como:

Dado $f \in V$, obtenha $g \in G$ que minimiza o erro quadrático (distância)

$$EQ(f, g) = \|f - g\| = \sqrt{\langle f - g, f - g\rangle}$$

entre f e os vetores de G.

MMQ Linear - Formulação Geral

Este problema admite uma única solução $\bar{g}$ que pode ser calculada da seguinte forma: escolha uma base $\{g_j\}_{j=0}^m$ de G. Então os coeficientes $\{a_j\}_{j=0}^m$ da representação

$$\bar{g} = \sum_{j=0}^{m} a_j g_j$$

são obtidos da resolução do **sistema normal**

$$\begin{pmatrix} \langle g_0, g_0 \rangle & \langle g_0, g_1 \rangle & \dots & \langle g_0, g_m \rangle \\ \langle g_1, g_0 \rangle & \langle g_1, g_1 \rangle & \dots & \langle g_1, g_m \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle g_m, g_0 \rangle & \langle g_m, g_1 \rangle & \dots & \langle g_m, g_m \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \langle g_0, f \rangle \\ \langle g_1, f \rangle \\ \vdots \\ \langle g_m, f \rangle \end{pmatrix}.$$

Bases ortogonais

Se a base de G for ortogonal,

$$\langle g_i, g_j \rangle = 0, \quad \text{se } i \neq j,$$

então o sistema normal terá matriz diagonal e será dado por

$$\langle g_i, g_i \rangle a_i = \langle g_i, f \rangle, \quad 0 \leq i \leq m.$$

A solução $\bar{g}$ do problema é então representada nesta base ortogonal por

$$\bar{g} = \sum_{j=0}^{m} \frac{\langle g_j, f \rangle}{\langle g_j, g_j \rangle} g_j. \tag{1}$$

Bases ortogonais

- Se a base escolhida para G não for ortogonal, é sempre possível a partir dela obter uma base ortogonal. Por exemplo, usando o processo de Gram-Schmidt;

Bases ortogonais

- Se a base escolhida para G não for ortogonal, é sempre possível a partir dela obter uma base ortogonal. Por exemplo, usando o processo de Gram-Schmidt;
- muitas vezes o interesse é no espaço G e não nas funções originalmente sugeridas para a aproximação. Por exemplo, aproximar por uma função da forma $g(x) = a_0 + a_1 x + a_2 x^2$ é equivalente a aproximar por um polinômio de grau menor ou igual a 2, e podemos usar qualquer base neste espaço;

Bases ortogonais

- Se a base escolhida para G não for ortogonal, é sempre possível a partir dela obter uma base ortogonal. Por exemplo, usando o processo de Gram-Schmidt;
- muitas vezes o interesse é no espaço G e não nas funções originalmente sugeridas para a aproximação. Por exemplo, aproximar por uma função da forma $g(x) = a_0 + a_1 x + a_2 x^2$ é equivalente a aproximar por um polinômio de grau menor ou igual a 2, e podemos usar qualquer base neste espaço;
- como veremos, há várias situações nas quais bases ortogonais já são conhecidas, permitindo-nos usá-las sem a necessidade de construí-las.

Bases ortogonais

Exemplo Considere o problema de se aproximar $f(x) = \sqrt{|x|}$ por um polinômio g de grau menor ou igual a 2, de forma a minimizar o erro quadrático

$$EQ(f,g) = \sqrt{\int_{-1}^{1} [f(x) - g(x)]^2 \, dx}.$$

Este erro quadrático vem do produto interno

$$\langle u, v \rangle = \int_{-1}^{1} u(x)v(x) \, dx$$

e sabe-se que os polinômios $p_0(x) = 1$, $p_1(x) = x$ e $p_2(x) = \frac{1}{2}(3x^2 - 1)$ satisfazem

$$\langle p_k, p_l \rangle = \begin{cases} 0 & \text{se } k \neq l, \\ \frac{2}{2k+1} & \text{se } k = l. \end{cases}$$

Bases ortogonais

Como p_0, p_1 e p_2 formam uma base para o espaço dos polinômios de grau menor ou igual a 2 (por que?), podemos buscar a aproximação na forma

$$g(x) = c_0 p_0(x) + c_1 p_1(x) + c_2 p_2(x).$$

Usando as relações de ortogonalidade do slide anterior e (1) obtemos

$$\begin{aligned}
c_0 &= \frac{\langle p_0, f\rangle}{\langle p_0, p_0\rangle} = \frac{1}{2}\int_{-1}^{1} |x|^{1/2}\, dx = \int_0^1 x^{1/2}\, dx = \frac{2}{3}, \\
c_1 &= \frac{\langle p_1, f\rangle}{\langle p_1, p_1\rangle} = \frac{1}{2/3}\int_{-1}^{1} x|x|^{1/2}\, dx = 0, \\
c_2 &= \frac{\langle p_2, f\rangle}{\langle p_2, p_2\rangle} = \frac{1}{2/5}\int_{-1}^{1} \frac{1}{2}(3x^2-1)|x|^{1/2}\, dx = \frac{5}{2}\int_0^1 (3x^{5/2} - x^{1/2}\, dx) \\
&= \frac{10}{21}.
\end{aligned}$$

Bases ortogonais

Protanto a aproximação fica

$$g(x) = \frac{2}{3} + \frac{10}{21}\frac{1}{2}(3x^2 - 1) = \frac{2}{3} + \frac{5}{21}(3x^2 - 1)$$
$$= \frac{3}{7} + \frac{5}{7}x^2.$$

Se buscássemos a aproximação na forma $g(x) = a_0 + a_1 x + a_2 x^2$, a solução do sitema normal formado a partir das funções $g_0(x) = 1$, $g_1(x) = x$ e $g_2(x) = x^2$ seria $a_0 = 3/7$, $a_1 = 0$ e $a_2 = 5/7$, pois a aproximação g é a mesma. $\square$

Polinômios ortogonais

Famílias ortogonais têm uma rica história na física-matemática, onde a solução de problemas pode envolver a aproximação ou expansão de funções por membros destas famílias. Famílias de polinômios ortogonais formam uma classe importante e serão objeto do nosso estudo. Primeiramente, fixemos a notação.

Notação

- O espaço vetorial dos polinômios com coeficientes reais e domínio $\mathbb{R}$ será denotado por $\mathcal{P}$;
- dado um número natural m, o subespaço vetorial de $\mathcal{P}$ formado pelos polinômios de grau menor ou igual a m será denotado por $\mathcal{P}_m$.

Observações:

1. A dimensão de $\mathcal{P}_m$ é $m + 1$;
2. um polinômio de *grau zero* é por definição uma função constante.

Polinômios ortogonais

Dado um produto interno $\langle\ ,\ \rangle$ em $\mathcal{P}$, uma família de polinômios ortogonais em relação a este produto interno é uma família $\{p_n\}$, $n = 0, 1, 2, \ldots,$ de elementos de $\mathcal{P}$ onde

1. o grau de p_n é igual a n;
2. $\langle p_n, p_m \rangle = 0$ se $n \neq m$.

Observações:

Polinômios ortogonais

Dado um produto interno $\langle\ ,\ \rangle$ em $\mathcal{P}$, uma família de polinômios ortogonais em relação a este produto interno é uma família $\{p_n\}$, $n = 0, 1, 2, \ldots$, de elementos de $\mathcal{P}$ onde

1. o grau de p_n é igual a n;
2. $\langle p_n, p_m \rangle = 0$ se $n \neq m$.

Observações:

- É possível gerar uma família de polinômios ortogonais aplicando-se o processo de Gram-Schmidt à base canônica $\{1, x, x^2, \ldots\}$ de $\mathcal{P}$;

Polinômios ortogonais

Dado um produto interno $\langle \ , \ \rangle$ em $\mathcal{P}$, uma família de polinômios ortogonais em relação a este produto interno é uma família $\{p_n\}$, $n = 0, 1, 2, \dots,$ de elementos de $\mathcal{P}$ onde

1. o grau de p_n é igual a n;
2. $\langle p_n, p_m \rangle = 0$ se $n \neq m$.

Observações:

- É possível gerar uma família de polinômios ortogonais aplicando-se o processo de Gram-Schmidt à base canônica $\{1, x, x^2, \dots\}$ de $\mathcal{P}$;
- há uma infinidade de famílias de polinômios ortogonais relativamente ao mesmo produto interno. De fato, se $\{p_n\}$ é uma família de polinômios ortogonais, então, escolhendo-se números reais arbitrários $\alpha_n \neq 0$, temos que $\{q_n\}$, onde $q_n(x) = \alpha_n p_n(x)$, é também uma família de polinômios ortogonais.

Polinômios ortogonais

Dado um produto interno $\langle\ ,\ \rangle$ em $\mathcal{P}$, uma família de polinômios ortogonais em relação a este produto interno é uma família $\{p_n\}$, $n = 0, 1, 2, \ldots$, de elementos de $\mathcal{P}$ onde

1. o grau de p_n é igual a n;
2. $\langle p_n, p_m \rangle = 0$ se $n \neq m$.

Observações:

- É possível gerar uma família de polinômios ortogonais aplicando-se o processo de Gram-Schmidt à base canônica $\{1, x, x^2, \ldots\}$ de $\mathcal{P}$;
- há uma infinidade de famílias de polinômios ortogonais relativamente ao mesmo produto interno. De fato, se $\{p_n\}$ é uma família de polinômios ortogonais, então, escolhendo-se números reais arbitrários $\alpha_n \neq 0$, temos que $\{q_n\}$, onde $q_n(x) = \alpha_n p_n(x)$, é também uma família de polinômios ortogonais.
- Se $\{p_n\}$, $n = 0, 1, \ldots$ é uma família de polinômios ortogonais, então $\{p_0, p_1, \ldots, p_m\}$ é uma base de $\mathcal{P}_m$.

Polinômios ortogonais

Veremos a seguir alguns exemplos classicos de famílias de polinômios ortogonais que aparecem em aplicações.

Exemplo 1

Polinômios de Legendre É a família formada pelos polinômios $\{p_n\}_{n=0}^{\infty}$ que satisfazem

$$\langle p_n, p_m \rangle = \int_{-1}^{1} p_n(x) p_m(x)\, dx = 0 \quad \text{se } n \neq m,$$

com a padronização $p_n(1) = 1$ para todo $n \geq 0$. Para estes polinômios temos

$$\langle p_n, p_n \rangle = \frac{2}{2n+1}.$$

Polinômios ortogonais

Exemplo 2

Polinômios de Chebyshev É a família $\{T_n\}_{n=0}^{\infty}$ definida por

$$\langle T_n, T_m \rangle = \int_{-1}^{1} \frac{T_n(x) T_m(x)}{\sqrt{1-x^2}}\, dx = 0 \quad \text{se } n \neq m,$$

com a padronização $T_n(1) = 1$ para todo $n \geq 0$. Note que o intervalo de integração é o mesmo que o dos polinômios de Legendre, mas temos a *função peso*

$$\omega(x) = \frac{1}{\sqrt{1-x^2}}.$$

Estes polinômios satisfazem $\langle T_0, T_0 \rangle = \pi$ e $\langle T_n, T_n \rangle = \frac{\pi}{2}$ se $n \geq 1$.

Polinômios ortogonais

Exemplo 3

Polinômios de Hermite É a família $\{H_n\}_{n=0}^{\infty}$ definida por

$$\langle H_n, H_m \rangle = \int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x)\, dx = 0 \quad \text{se } n \neq m,$$

e tal que o coeficiente de x^n de H_n é igual 2^n, $n \geq 0$. O intervalo de integração é $(-\infty, \infty)$ e temos a função peso

$$\omega(x) = e^{-x^2}.$$

Estes polinômios satisfazem $\langle H_n, H_n \rangle = \sqrt{\pi} 2^n n!$.

Polinômios ortogonais

Exemplo 4

Polinômios de Laguerre É a família $\{L_n\}_{n=0}^{\infty}$ definida por

$$\langle L_n, L_m \rangle = \int_0^{\infty} e^{-x} L_n(x) L_m(x)\, dx = 0 \quad \text{se } n \neq m,$$

e tal que o coeficiente de x^n de L_n é igual a $\frac{(-1)^n}{n!}$, $n \geq 0$. Agora o intervalo de integração é $[0, \infty)$, com a função peso dada por

$$\omega(x) = e^{-x}.$$

Estes polinômios satisfazem $\langle L_n, L_n \rangle = 1$, e portanto são ortonormais.

Mudança de variável

Esses exemplos são de famílias bem estudadas, cujos membros podem ser encontrados em tabelas ou gerados por relações de recorrências bem conhecidas. Este conhecimento pode ser aproveitado no método dos mínimos quadrados, por exemplo, da seguinte forma.

Considere o problema de se aproximar uma função $f : [a, b] \to \mathbb{R}$ por um polinômio g de grau menor ou igual a m, de forma a minimizar $\int_a^b [f(x) - g(x)]^2 \, dx$. Este é um problema de mínimos quadrados onde o produto interno é $\langle u, v \rangle = \int_a^b u(x)v(x) \, dx$. Sabemos que os polinômios de Legendre são ortogonais em $[-1, 1]$. Como usar esta informação? A idéia é fazer uma mudança de variável por uma transformação linear afim, conforme descrito a seguir.

Mudança de variável

Do cálculo, sabemos que, se fizermos a mudança de variável $\varphi : [-1, 1] \to [a, b]$ definida por $\varphi(t) = \frac{b-a}{2} t + \frac{a+b}{2}$, então

$$\int_a^b [f(x) - g(x)]^2 \, dx = \frac{b-a}{2} \int_{-1}^{1} [f(\varphi(t)) - g(\varphi(t))]^2 \, dt. \qquad (2)$$

Tranformações lineares afins preservam graus de polinômios. De (2), concluímos então que aproximações polinomiais em $[a, b]$ e $[-1, 1]$ que minimizam os respectivos erros quadráticos relacionam-se por transformações lineares afins. O problema pode ser resolvido da seguinte forma:

Mudança de variável

1. aproxime $F(t) = f(\varphi(t))$ por um polinômio G de grau de grau menor ou igual a m de forma a minimizar $\int_{-1}^{1} [F(t) - G(t)]^2 \, dt$;

Mudança de variável

1. aproxime $F(t) = f(\varphi(t))$ por um polinômio G de grau de grau menor ou igual a m de forma a minimizar $\int_{-1}^{1} [F(t) - G(t)]^2 \, dt$;
2. usando os polinômios de Legendre $\{p_k\}_{k=0}^{m}$, esta aproximação é $G(t) = \sum_{k=0}^{m} c_k p_k(t)$, com

$$c_k = \frac{(p_k|F)}{(p_k|p_k)} = \frac{2k+1}{2} \int_{-1}^{1} p_k(t) F(t)$$

segundo (1) com o produto interno $(u|v) = \int_{-1}^{1} u(t)v(t) \, dt$;

Mudança de variável

1. aproxime $F(t) = f(\varphi(t))$ por um polinômio G de grau de grau menor ou igual a m de forma a minimizar $\int_{-1}^{1} [F(t) - G(t)]^2 \, dt$;
2. usando os polinômios de Legendre $\{p_k\}_{k=0}^{m}$, esta aproximação é $G(t) = \sum_{k=0}^{m} c_k p_k(t)$, com

$$c_k = \frac{(p_k|F)}{(p_k|p_k)} = \frac{2k+1}{2} \int_{-1}^{1} p_k(t) F(t)$$

   segundo (1) com o produto interno $(u|v) = \int_{-1}^{1} u(t)v(t) \, dt$;
3. a solução g do problema é obtida usando-se a transformação linear afim inversa $\psi : [a, b] \to [-1, 1]$ dada por $\psi(x) = \frac{2}{b-a} x - \frac{a+b}{b-a}$:

$$g(x) = G(\psi(x)) = \sum_{k=0}^{m} c_k p_k(\psi(x)).$$

Mudança de variável

Exemplo 5

Aproxime a função $f(x) = 1 - |2x - 5|$ no intervalo $[2, 3]$ por um polinômio g de grau menor ou igual a 3 de forma a minimizar $\int_2^3 [f(x) - g(x)]^2\, dx$. Sabe-se que os polinômios de Legendre $p_0(t) = 1$, $p_1(t) = t$, $p_2(t) = \frac{1}{2}(3t^2 - 1)$ e $p_3(t) = \frac{1}{2}(5t^3 - 3t)$ satisfazem

$$(p_n|p_m) = \int_{-1}^{1} p_n(x)p_m(x)\, dx = \begin{cases} 0 & \text{se } n \neq m \\ \frac{2}{2n+1} & \text{se } n = m \end{cases}$$

Este é um problema de MMQ cujo erro quadrático está associado ao produto interno $\langle u, v\rangle = \int_2^3 u(x)v(x)\, dx$. Podemos usar mudança de variável para explorar a ortogonalidade dos polinômios de Legendre no intervalo $[-1, 1]$.

Mudança de variável

A transformação $\varphi(t) = \frac{1}{2}t + \frac{5}{2}$ leva o intervalo $[-1, 1]$ no intervalo $[2, 3]$. Logo, primeiramente vamos aproximar $F(t) = f(\varphi(t)) = 1 - |t|$ por $G(t) = \sum_{k=0}^{3} c_k p_k(t)$ no intervalo $[-1, 1]$. Os coeficientes c_k são dados por

$$c_0 = \frac{(p_0|F)}{(p_0|p_0)} = \frac{1}{2}\int_{-1}^{1}(1 - |t|)\,dt = \int_{0}^{1}(1 - t)\,dt = \frac{1}{2}$$

$$c_1 = \frac{(p_1|F)}{(p_1|p_1)} = \frac{3}{2}\int_{-1}^{1} t(1 - |t|)\,dt = 0$$

$$c_2 = \frac{(p_2|F)}{(p_2|p_2)} = \frac{5}{2}\int_{-1}^{1}\frac{1}{2}(3t^2 - 1)(1 - |t|)\,dt = \frac{5}{2}\int_{0}^{1}(3t^2 - 1)(1 - t)\,dt$$
$$= -\frac{1}{4}$$

$$c_3 = \frac{(p_3|F)}{(p_3|p_3)} = \frac{7}{2}\int_{-1}^{1}\frac{1}{2}(5t^3 - 3t)(1 - |t|)\,dt = 0$$

Mudança de variável

A aproximação em $[-1, 1]$ fica

$$G(t) = \frac{1}{2}p_0(t) - \frac{1}{4}p_2(t) = \frac{1}{2} - \frac{1}{8}(3t^2 - 1).$$

Como a transformação inversa de $[2, 3]$ em $[-1, 1]$ é $\psi(x) = 2x - 5$, a solução do nosso exemplo é dada por

$$g(x) = G(\psi(x)) = \frac{1}{2} - -\frac{1}{8}[3(2x - 5)^2 - 1].$$

$\square$