Método dos Mínimos Quadrados (MMQ)

Produto Interno

Nelson Kuhl

IME/USP

15 de setembro de 2020

Introdução

O conceito matemático mais importante para o problema de mínimos quadrados linear é o de um *espaço vetorial munido de um produto interno*. Ele permite abstrair as propriedades essenciais do $\mathbb{R}^n$ com o seu produto escalar que nos levaram à solução de um sistema linear sobredeterminado no sentido de mínimos quadrados. Em um espaço vetorial com um produto interno, podemos definir comprimento de vetores, distância entre vetores e, o mais importante, **a noção de ortogonalidade**. Restringiremo-nos apenas ao caso de espaços vetoriais reais. O caso em que os escalares são complexos é importante mas está além dos nossos objetivos.

Espaço Vetorial com Produto Interno

Seja V um espaço vetorial real. Um produto interno em V é uma aplicação

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

satisfazendo:

Espaço Vetorial com Produto Interno

Seja V um espaço vetorial real. Um produto interno em V é uma aplicação

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

satisfazendo:

(P1) $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$;

Espaço Vetorial com Produto Interno

Seja V um espaço vetorial real. Um produto interno em V é uma aplicação

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

satisfazendo:

(P1) $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$;

(P2) $\langle u, v\rangle = \langle v, u\rangle$, $\forall u, v \in V$;

Espaço Vetorial com Produto Interno

Seja V um espaço vetorial real. Um produto interno em V é uma aplicação

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

satisfazendo:

(P1) $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$;

(P2) $\langle u, v\rangle = \langle v, u\rangle$, $\forall u, v \in V$;

(P3) $\langle u, u\rangle > 0$, $\forall u \in V$, $u \neq 0$.

Espaço Vetorial com Produto Interno

Seja V um espaço vetorial real. Um produto interno em V é uma aplicação

$$\langle \ , \ \rangle : V \times V \to \mathbb{R}$$

satisfazendo:

(P1) $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$;

(P2) $\langle u, v\rangle = \langle v, u\rangle$, $\forall u, v \in V$;

(P3) $\langle u, u\rangle > 0$, $\forall u \in V$, $u \neq 0$.

Observação: Note que (P1) implica $\langle u, u\rangle = 0$ quando $u = 0$.

Espaço Vetorial com Produto Interno

Espaço Vetorial com Produto Interno

- Usando-se a linearidade em relação à segunda componente (P1) e a simetria (P2), concluimos que $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$. Portanto o produto interno é *bilinear*;

Espaço Vetorial com Produto Interno

- Usando-se a linearidade em relação à segunda componente (P1) e a simetria (P2), concluimos que $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$. Portanto o produto interno é *bilinear*;
- a propriedade (P3) afirma que o produto interno é *positivo definido*;

Espaço Vetorial com Produto Interno

- Usando-se a linearidade em relação à segunda componente (P1) e a simetria (P2), concluimos que $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$. Portanto o produto interno é *bilinear*;
- a propriedade (P3) afirma que o produto interno é *positivo definido*;
- de forma sucinta, podemos definir um produto interno em um espaço vetorial real V como **uma forma bilinear, simétrica e definida positiva** em V;

Espaço Vetorial com Produto Interno

- Usando-se a linearidade em relação à segunda componente (P1) e a simetria (P2), concluimos que $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$. Portanto o produto interno é *bilinear*;
- a propriedade (P3) afirma que o produto interno é *positivo definido*;
- de forma sucinta, podemos definir um produto interno em um espaço vetorial real V como **uma forma bilinear, simétrica e definida positiva** em V;
- o **comprimento** de um vetor u em um espaço V munido de um produto interno é definido por $\|u\| = \sqrt{\langle u, u\rangle}$, e a **distância** $\|u - v\|$ entre dois vetores u e v de V é o comprimento de $u - v$;

Espaço Vetorial com Produto Interno

- Usando-se a linearidade em relação à segunda componente (P1) e a simetria (P2), concluimos que $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v \in V$. Portanto o produto interno é *bilinear*;
- a propriedade (P3) afirma que o produto interno é *positivo definido*;
- de forma sucinta, podemos definir um produto interno em um espaço vetorial real V como **uma forma bilinear, simétrica e definida positiva** em V;
- o **comprimento** de um vetor u em um espaço V munido de um produto interno é definido por $\|u\| = \sqrt{\langle u, u\rangle}$, e a **distância** $\|u - v\|$ entre dois vetores u e v de V é o comprimento de $u - v$;
- de (P1) e (P3) segue que **somente o vetor nulo tem comprimento zero** e protanto se a distância entre dois vetores for zero, eles são idênticos.

Exemplos

Exemplo 1 Em $V = \mathbb{R}^n$, o produto escalar

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

é um produto interno.

Exemplo 2 Uma generalização do produto escalar é obtida da seguinte forma: considere n números reais *positivos* dados $\{\omega_i\}_{i=1}^{n}$. Então,

$$\langle x, y \rangle = \sum_{i=1}^{n} \omega_i x_i y_i$$

é um produto interno no $\mathbb{R}^n$. Com este produto interno, a distância entre dois vetores x e y do $\mathbb{R}^n$ fica

$$\|x - y\| = \sqrt{\sum_{i=1}^{n} \omega_i (x_i - y_i)^2}.$$

Exemplos

Exemplo 3 Seja $V = C([a,b])$ o espaço vetorial das funções reais contínuas definidas no intervalo $[a,b]$. Então

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$$

é um produto interno em V e a distância entre duas funções é igual a

$$\|f - g\| = \sqrt{\int_a^b [f(x) - g(x)]^2\,dx}.$$

Exemplos

Exemplo 4 Generalizando o exemplo anterior, considere uma função dada $\omega : (a, b) \to \mathbb{R}$ contínua, positiva e tal que $\int_a^b \omega(x)\, dx$ exista. Então

$$\langle f, g\rangle = \int_a^b \omega(x) f(x) g(x)\, dx \tag{1}$$

define um produto interno em $C([a, b])$ e a distância entre duas funções f e g, segundo este produto interno, é igual a

$$\|f - g\| = \sqrt{\int_a^b \omega(x)[f(x) - g(x)]^2\, dx}. \tag{2}$$

Exemplos

Exemplo 5 Seja $\mathcal{P}$ o espaço vetorial formado pelos polinômios com coeficientes reais definidos em $\mathbb{R}$. Dados os números reais a e b, com $a < b$, a expressão (1) define um produto interno em $\mathcal{P}$, com a distância entre dois polinômios dada por (2).

Exemplos

Exemplo 5 Seja $\mathcal{P}$ o espaço vetorial formado pelos polinômios com coeficientes reais definidos em $\mathbb{R}$. Dados os números reais a e b, com $a < b$, a expressão (1) define um produto interno em $\mathcal{P}$, com a distância entre dois polinômios dada por (2).

O exemplo a seguir é muito importante, pois enfraquece a noção de produto interno e permite a formalização do caso de domínios discretos para o MMQ.

Exemplos

Exemplo 6 Seja V um espaço vetoral formado por funções reais possuindo um domínio comum. Sejam $\{x_i\}_{i=1}^n$ pontos distintos pertencentes ao domínio das funções e $\{\omega_i\}_{i=1}^n$ números positivos dados. Para f e g pertencentes a V, defina a forma

$$\langle f, g \rangle = \sum_{i=1}^{n} \omega_i f(x_i) g(x_i). \tag{3}$$

Esta forma satisfaz (P1) e (P2) mas o máximo que podemos garantir é que ela é **semidefinida positiva**, $\langle f, f \rangle \geq 0$, pois podemos ter $f(x_i) = 0$, $1 \leq i \leq n$, sem que f seja identicamente nula. Mesmo assim definimos a distância entre f e g por

$$\|f - g\| = \sqrt{\langle f - g, f - g \rangle} = \sqrt{\sum_{i=1}^{n} \omega_i [f(x_i) - g(x_i)]^2}. \tag{4}$$

Produtos internos degenerados

Mesmo não sendo uma terminologia padrão, daremos a seguinte definição devido à sua importância para o MMQ.

Definição. Dado um espaço vetorial real V, um *produto interno degenerado* em V é uma aplicação $\langle \ , \ \rangle : V \times V \to \mathbb{R}$ satisfazendo (P1), (P2) e

$$\text{(P3')} \qquad \langle u, u \rangle \geq 0, \quad \forall u \in V.$$

Ou seja, uma forma *bilinear, simétrica e semidefinida positiva*.

Produtos internos degenerados

Mesmo não sendo uma terminologia padrão, daremos a seguinte definição devido à sua importância para o MMQ.

Definição. Dado um espaço vetorial real V, um *produto interno degenerado* em V é uma aplicação $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ satisfazendo (P1), (P2) e

$$\text{(P3')} \qquad \langle u, u \rangle \geq 0, \quad \forall u \in V.$$

Ou seja, uma forma *bilinear, simétrica e semidefinida positiva*.

- O comprimento de um vetor e a distância entre vetores são definidos como no caso de um produto interno;

Produtos internos degenerados

Mesmo não sendo uma terminologia padrão, daremos a seguinte definição devido à sua importância para o MMQ.

Definição. Dado um espaço vetorial real V, um *produto interno degenerado* em V é uma aplicação $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ satisfazendo (P1), (P2) e

$$\text{(P3')} \qquad \langle u, u \rangle \geq 0, \quad \forall u \in V.$$

Ou seja, uma forma *bilinear, simétrica e semidefinida positiva*.

- O comprimento de um vetor e a distância entre vetores são definidos como no caso de um produto interno;
- a distância $\|u - v\| = \sqrt{\langle u - v, u - v \rangle}$ entre dois vetores será chamada também de **erro quadrático** entre eles;

Produtos internos degenerados

Mesmo não sendo uma terminologia padrão, daremos a seguinte definição devido à sua importância para o MMQ.

Definição. Dado um espaço vetorial real V, um *produto interno degenerado* em V é uma aplicação $\langle \ , \ \rangle : V \times V \to \mathbb{R}$ satisfazendo (P1), (P2) e

$$\text{(P3')} \qquad \langle u, u \rangle \geq 0, \quad \forall u \in V.$$

Ou seja, uma forma *bilinear, simétrica e semidefinida positiva*.

- O comprimento de um vetor e a distância entre vetores são definidos como no caso de um produto interno;
- a distância $\|u - v\| = \sqrt{\langle u - v, u - v \rangle}$ entre dois vetores será chamada também de **erro quadrático** entre eles;
- o vetor nulo tem comprimento zero, mas podemos ter vetores não nulos com comprimento zero e vetores distintos cuja distância entre eles é zero.

Ortogonalidade

Seja V um espaço vetorial real com um produto interno degenerado $\langle \ , \ \rangle$.

Ortogonalidade

Seja V um espaço vetorial real com um produto interno degenerado $\langle\ ,\ \rangle$.

- Dizemos que dois vetores $u, v \in V$ são ortogonais, $u \perp v$, se $\langle u, v\rangle = 0$;

Ortogonalidade

Seja V um espaço vetorial real com um produto interno degenerado $\langle \;,\; \rangle$.

- Dizemos que dois vetores $u, v \in V$ são ortogonais, $u \perp v$, se $\langle u, v \rangle = 0$;
- dizemos que $u \in V$ é ortogonal a um subespaço vetorial $G \subset V$, $u \perp G$, se $u \perp g$ para todo $g \in G$;

Ortogonalidade

Seja V um espaço vetorial real com um produto interno degenerado $\langle\ ,\ \rangle$.

- Dizemos que dois vetores $u, v \in V$ são ortogonais, $u \perp v$, se $\langle u, v\rangle = 0$;
- dizemos que $u \in V$ é ortogonal a um subespaço vetorial $G \subset V$, $u \perp G$, se $u \perp g$ para todo $g \in G$;
- dados $u \in V$ e $G \subset V$ subespaço vetorial, dizemos que $\bar{g} \in G$ é uma projeção ortogonal de u em G se $u - \bar{g} \perp G$.

Projeções ortogonais e distância mínima

Lema 1

Sejam V um espaço vetorial real munido de um produto interno degenerado $\langle\ ,\ \rangle$, $f \in V$ e $G \subset V$ um subespaço vetorial. Se $\bar{g} \in G$ é uma projeção ortogonal de f em G, então

$$\|f - \bar{g}\| \leq \|f - g\|, \quad \forall g \in G.$$

Ou seja, $\bar{g}$ minimiza o erro quadrático entre f e os elementos de G.

Projeções ortogonais e distância mínima

Demonstração

Para $g \in G$ temos

$$\begin{aligned}\langle f-g, f-g\rangle &= \langle f-\bar{g}+\bar{g}-g, f-\bar{g}+\bar{g}-g\rangle \\ &= \langle f-\bar{g}, f-\bar{g}\rangle + 2\langle f-\bar{g}, \bar{g}-g\rangle + \langle \bar{g}-g, \bar{g}-g\rangle\end{aligned}$$

onde na última igualdade usamos a bilinearidade e a simetria. Como $f-\bar{g} \perp G$ e $\bar{g}-g \in G$, temos $\langle f-\bar{g}, \bar{g}-g\rangle = 0$, e como $\langle \bar{g}-g, \bar{g}-g\rangle \geq 0$, concluimos que

$$\|f-g\|^2 \geq \|f-\bar{g}\|^2, \quad \forall g \in G,$$

o que nos dá o resultado.

□

Observações

Observações

- Note que usamos as propriedades (P1), (P2) e (P3'), abstraídas do produto escalar;

Observações

- Note que usamos as propriedades (P1), (P2) e (P3'), abstraídas do produto escalar;
- assumimos a existência da projeção ortogonal, sem demonstrá-la;

Observações

- Note que usamos as propriedades (P1), (P2) e (P3'), abstraídas do produto escalar;
- assumimos a existência da projeção ortogonal, sem demonstrá-la;
- iremos demonstrar a existência da projeção ortogonal no caso em que G tem dimensão finita e a *restrição* de $\langle\ ,\ \rangle$ a G é um *produto interno* em G, ou seja, satisfaz (P1), (P2) e **(P3)** em G.

Existência da projeção ortogonal

Teorema 1

Sejam V um espaço vetorial real munido de um produto interno semidefinido $\langle\ ,\ \rangle$, $f \in V$ e $G \subset V$ um subespaço vetorial de *dimensão finita* tal que a restrição de $\langle\ ,\ \rangle$ a ele é um produto interno. Então existe um único $\bar{g} \in G$ tal que $f - \bar{g} \perp G$.

Demonstração

Suponha que a dimensão de G é $m+1$ (para ficar de acordo com várias aplicações). Seja $\{g_j\}_{j=0}^m$ uma *base* de G. Queremos então mostrar que existe um único $\bar{g} \in G$ tal que

$$f - \bar{g} \perp g_i, 0 \leq i \leq m,$$

ou, expandindo-se $\bar{g}$ na base, $\bar{g} = \sum_{j=0}^m a_j g_j$, queremos mostrar que existem únicos coeficientes $\{a_j\}_{j=0}^m$ tais que

Existência da projeção ortogonal

$$\langle f - \sum_{j=0}^{m} a_j g_j, g_i \rangle = 0, \quad 0 \le i \le m.$$

Usando a simetria e a bilinearidade, podemos expandir as expressões acima e obter o sitema linear

$$\sum_{j=0}^{m} \langle g_i, g_j \rangle a_j = \langle g_i, f \rangle, \quad 0 \le i \le m, \tag{5}$$

para os coeficientes $\{a_j\}_{j=0}^{m}$, conhecido como (o já famoso) **sistema normal**. Sendo $\{g_j\}_{j=0}^{m}$ uma família linearmente independente e $\langle \ , \ \rangle$ um produto interno em G, a matriz de (5) é inversível e portanto sistema normal tem solução única.

$\square$

Método dos Mínimos Quadrados Linear

Usando o Lema 1 e parte da demonstração do Teorema 1, podemos demonstrar o seguinte resultado, que é uma formulação abstrata do Método dos Mínimos Quadrados.

Teorema 2

Sejam V um espaço vetorial real munido de um produto interno degenerado $\langle\ ,\ \rangle$, $f \in V$ e $G \subset V$ um subespaço vetorial de dimensão finita $m+1$. Seja $\{g_j\}_{j=0}^m$ uma base de G. Se o sitema normal (5) tiver solução $\{a_j\}_{j=0}^m$, então

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

minimiza o erro quadrático entre f e os vetores de G, ou seja, $\|f - \bar{g}\| \leq \|f - g\|$, $\forall g \in G$. Se a matriz do sistema normal (5) for inversível, então existe um único $\bar{g} \in G$ que minimiza o erro quadrático entre f e os vetores de G.

Método dos Mínimos Quadrados Linear

Frequentemente em aplicações, queremos aproximar uma função $f(x)$ por uma combinação linear

$$g(x) = \sum_{j=0}^{m} a_j g_j(x)$$

de funções $\{g_j(x)\}_{j=0}^{m}$ especificadas, segundo um critério de erro. Quando este erro pode ser associado à distância definida por um produto interno, temos um problema de mínimos quadrados linear, onde V é um espaço vetorial contendo $\{f, g_0, \ldots, g_m\}$ e G é o subespaço vetorial gerado por $\{g_j\}_{j=0}^{m}$. Veremos adiante uma série de exemplos e aplicações. O formalismo é sempre o mesmo, mudando-se os produtos internos e espaços vetoriais.