O Método dos Mínimos Quadrados

Prof. Afonso Paiva
Prof. Fabricio Simeoni

Departamento de Matemática Aplicada e Estatística
Instituto de Ciências Matemáticas e de Computação
USP – São Carlos

Cálculo Numérico – SME0300

Método dos mínimos quadrados

- O objetivo do método é aproximar uma função qualquer (conhecida ou não) por uma combinação de funções conhecidas;

Método dos mínimos quadrados

- O objetivo do método é aproximar uma função qualquer (conhecida ou não) por uma combinação de funções conhecidas;
- Garantias teóricas: aproximação é a melhor possível;

Método dos mínimos quadrados

- O objetivo do método é aproximar uma função qualquer (conhecida ou não) por uma combinação de funções conhecidas;
- Garantias teóricas: aproximação é a melhor possível;
- Praticidade: resultado de fácil obtenção e manipulação;

Motivação

Os motivos são os mais variados. Por exemplo, pode-se querer manipular uma função complicada, como

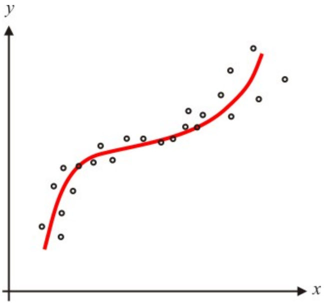
$$f(x) = e^{\sin(\cos(\sinh(\cosh(\tan^{-1}(\log(x))))))}$$

Motivação

Os motivos são os mais variados. Por exemplo, pode-se querer manipular uma função complicada, como

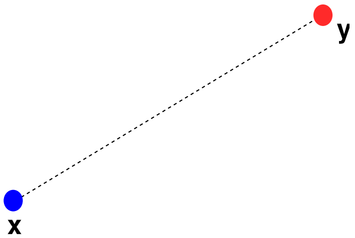
$$f(x) = e^{\sin(\cos(\sinh(\cosh(\tan^{-1}(\log(x))))))}$$

ou então encontrar uma aproximação para funções que nem são conhecidas, como por exemplo,



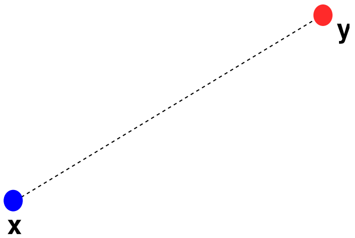
Distância em $\mathbb{R}^n$

Sejam $\mathbf{x} = (x_1, x_2)^\top$, $\mathbf{y} = (y_1, y_2)^\top \in \mathbb{R}^2$ dois pontos. Como medir a distância entre eles?



Distância em $\mathbb{R}^n$

Sejam $\mathbf{x} = (x_1, x_2)^\top$, $\mathbf{y} = (y_1, y_2)^\top \in \mathbb{R}^2$ dois pontos. Como medir a distância entre eles?

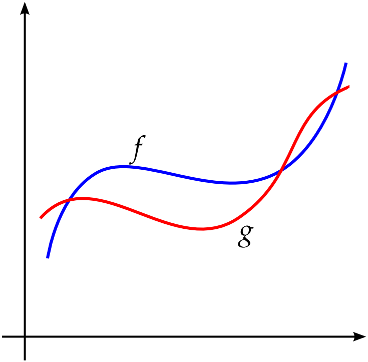


Distância Euclidiana:

$$dist(\mathbf{x}, \mathbf{y}) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2}$$

Distância em espaços de funções

Dadas funções $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$, como medir o quão próximo f está de g?



Precisamos então definir distâncias e construir a teoria necessária para escolher a melhor aproximação para uma função.

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- **associatividade**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *e* $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- **associatividade**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *e* $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;
- **elemento neutro**: *existe um vetor* $\bar{0} \in V$, *tal que* $\mathbf{v} + \bar{0} = \bar{0} + \mathbf{v} = \mathbf{v}$ *para todo* $\mathbf{v} \in V$;

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- **associatividade**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *e* $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;
- **elemento neutro**: *existe um vetor* $\bar{0} \in V$, *tal que* $\mathbf{v} + \bar{0} = \bar{0} + \mathbf{v} = \mathbf{v}$ *para todo* $\mathbf{v} \in V$;
- **inverso aditivo**: *para cada vetor* $\mathbf{v} \in V$ *existe um vetor* $-\mathbf{v} \in V$, *tal que* $-\mathbf{v} + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \bar{0}$;

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- **associatividade**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *e* $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;
- **elemento neutro**: *existe um vetor* $\bar{0} \in V$, *tal que* $\mathbf{v} + \bar{0} = \bar{0} + \mathbf{v} = \mathbf{v}$ *para todo* $\mathbf{v} \in V$;
- **inverso aditivo**: *para cada vetor* $\mathbf{v} \in V$ *existe um vetor* $-\mathbf{v} \in V$, *tal que* $-\mathbf{v} + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \bar{0}$;
- **distributiva**: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ *e* $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$;

Um pouco de álgebra linear

Definição (Espaço Vetorial)

Um espaço vetorial V é um conjunto que possui definidas as operações de soma e de produto (multiplicação por escalar pertencente a $\mathbb{R}$) e fechado com respeito a elas. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, as seguintes propriedades:

- **comutatividade**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- **associatividade**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *e* $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$;
- **elemento neutro**: *existe um vetor* $\bar{0} \in V$, *tal que* $\mathbf{v} + \bar{0} = \bar{0} + \mathbf{v} = \mathbf{v}$ *para todo* $\mathbf{v} \in V$;
- **inverso aditivo**: *para cada vetor* $\mathbf{v} \in V$ *existe um vetor* $-\mathbf{v} \in V$, *tal que* $-\mathbf{v} + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \bar{0}$;
- **distributiva**: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ *e* $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$;
- **multiplicação por 1**: $1 \cdot \mathbf{v} = \mathbf{v}$.

Um pouco de álgebra linear

Definição (Subespaço Vetorial)

Um subespaço vetorial de V é um subconjunto $S \subset V$ que satisfaz as seguintes propriedades:

- $\overline{0} \in S$;
- *se* $\mathbf{u}, \mathbf{v} \in S$ *então* $\mathbf{u} + \mathbf{v} \in S$;
- *se* $\mathbf{v} \in S$ *então* $\alpha\mathbf{v} \in S$, *para todo* $\alpha \in \mathbb{R}$.

Um pouco de álgebra linear

Definição (Subespaço Vetorial)

Um subespaço vetorial de V é um subconjunto $S \subset V$ que satisfaz as seguintes propriedades:

- $\overline{0} \in S$;
- *se* $\mathbf{u}, \mathbf{v} \in S$ *então* $\mathbf{u} + \mathbf{v} \in S$;
- *se* $\mathbf{v} \in S$ *então* $\alpha\mathbf{v} \in S$*, para todo* $\alpha \in \mathbb{R}$.

Teorema

Todo subespaço vetorial é um espaço vetorial.

Um pouco de álgebra linear

Exemplo 1

$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ é um espaço vetorial. Os hiperplanos de $\mathbb{R}^n$ que passam pela origem são subespaços vetoriais de $\mathbb{R}^n$.

Um pouco de álgebra linear

Exemplo 1

$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ é um espaço vetorial. Os hiperplanos de $\mathbb{R}^n$ que passam pela origem são subespaços vetoriais de $\mathbb{R}^n$.

Exemplo 2

O conjunto $M(n, n)$ das matrizes reais quadradas de ordem n é um espaço vetorial. O conjunto das matrizes simétricas de ordem n é um subspaço vetorial de $M(n, n)$.

Um pouco de álgebra linear

Exemplo 1

$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ é um espaço vetorial. Os hiperplanos de $\mathbb{R}^n$ que passam pela origem são subespaços vetoriais de $\mathbb{R}^n$.

Exemplo 2

O conjunto $M(n, n)$ das matrizes reais quadradas de ordem n é um espaço vetorial. O conjunto das matrizes simétricas de ordem n é um subspaço vetorial de $M(n, n)$.

Exemplo 3

Seja $\mathcal{F}(\mathbb{R}; \mathbb{R})$ o conjunto de todas as funções $f : \mathbb{R} \to \mathbb{R}$. Então são subspaços de $\mathcal{F}(\mathbb{R}; \mathbb{R})$:

$$\mathcal{P}_n \subset \mathcal{P} \subset \mathcal{C}^\infty(\mathbb{R}) \subset \mathcal{C}^k(\mathbb{R}) \subset \mathcal{C}^0(\mathbb{R})$$

Um pouco de álgebra linear

Definição (Conjunto L.I.)

Um conjunto $\mathcal{B} = \{\varphi_1, \varphi_2, \ldots \varphi_n\} \subset V$ *é dito linearmente independente (L.I.) se*

$$\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_n\varphi_n = \bar{0} \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_n = 0\,.$$

Um pouco de álgebra linear

Definição (Conjunto L.I.)

Um conjunto $\mathcal{B} = \{\varphi_1, \varphi_2, \ldots \varphi_n\} \subset V$ *é dito linearmente independente (L.I.) se*

$$\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_n\varphi_n = \bar{0} \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_n = 0\,.$$

Definição (Base de espaço vetorial)

Um conjunto $\mathcal{B} = \{\varphi_1, \ldots, \varphi_n\} \subset V$ *é uma base de um espaço vetorial* V *se for L.I. e gerar* V*. Isto é, todo vetor* $\mathbf{v} \in V$ *é escrito, de forma única, como combinação linear dos elementos de* $\mathcal{B}$*:*

$$\mathbf{v} = \alpha_1\varphi_1 + \ldots + \alpha_n\varphi_n\,.$$

Um pouco de álgebra linear

Definição (Dimensão)

A dimensão de um espaço vetorial V*, denotada por* $dim(V)$*, é o número máximo de elementos L.I. nele contido.*

Um pouco de álgebra linear

Definição (Dimensão)

A dimensão de um espaço vetorial V, denotada por $\dim(V)$, é o número máximo de elementos L.I. nele contido.

Teorema

Todo espaço vetorial de dimensão $n < \infty$ tem uma base.

Um pouco de álgebra linear

Definição (Dimensão)

A dimensão de um espaço vetorial V, denotada por $dim(V)$, é o número máximo de elementos L.I. nele contido.

Teorema

Todo espaço vetorial de dimensão $n < \infty$ tem uma base.

Definição

Se para qualquer conjunto de vetores de V sempre é possível encontrar um vetor L.I. à este conjunto, então dizemos que $dim(V) = \infty$.

Um pouco de álgebra linear

Exemplo 4

Seja $V = \mathbb{R}^n$. Uma base para V é o conjunto $\mathcal{B} = \{e_1, e_2, \ldots, e_n\} \subset \mathbb{R}^n$, onde

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i$$

Ainda, $dim(\mathbb{R}^n) = n$.

Um pouco de álgebra linear

Exemplo 5

Seja $V = M(m,n)$. Uma base para V é o conjunto $\mathcal{B} = \{E_{11}, \ldots, E_{mn}\} \subset M(m,n)$, onde:

$$E_{ij} = \begin{array}{cc} \begin{bmatrix} 0 & & & 0 & & & 0 \\ \vdots & & & \vdots & & & \vdots \\ \vdots & & & 0 & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & 0 & & & \vdots \\ \vdots & & & \vdots & & & \vdots \\ 0 & & & 0 & & & 0 \end{bmatrix} & \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow \\ j \end{matrix} & \end{array}$$

Ainda, $dim(M(m,n)) = mn$.

Um pouco de álgebra linear

Exemplo 6

Se $V = \mathcal{C}(\mathbb{R})$, então $dim(V) = \infty$.
Se $S = \mathcal{P}_n(\mathbb{R}) \subset V$, então $dim(S) = n + 1$. Uma base para S seria

$$\mathcal{B} = \{x^i : i = 0, \ldots, n\} = \{1, x, x^2, \ldots, x^n\}$$

É fácil verificar que todo polinômio $p \in \mathcal{P}_n$, de grau $\leq n$, pode ser escrito como

$$p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n$$

que é uma combinação linear dos elementos de $\mathcal{B}$.

Um pouco de álgebra linear

Definição (Produto Interno)

Uma aplicação $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ *é um produto interno se satisfaz as seguintes propriedades:*

- *bilinearidade:*

$$\begin{aligned} \langle \mathbf{v}, \alpha\mathbf{w} + \beta\mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{z} \rangle \\ \langle \alpha\mathbf{v} + \beta\mathbf{w}, \mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{z} \rangle + \beta\langle \mathbf{w}, \mathbf{z} \rangle, \end{aligned}$$

  $\forall\, \alpha, \beta \in \mathbb{R}, \forall\, \mathbf{v}, \mathbf{w}, \mathbf{z} \in V;$

Um pouco de álgebra linear

Definição (Produto Interno)

Uma aplicação $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ *é um produto interno se satisfaz as seguintes propriedades:*

- *bilinearidade:*

$$\begin{aligned} \langle \mathbf{v}, \alpha\mathbf{w} + \beta\mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{z} \rangle \\ \langle \alpha\mathbf{v} + \beta\mathbf{w}, \mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{z} \rangle + \beta\langle \mathbf{w}, \mathbf{z} \rangle, \end{aligned}$$

  $\forall\, \alpha, \beta \in \mathbb{R}, \forall\, \mathbf{v}, \mathbf{w}, \mathbf{z} \in V$;

- *comutatividade (simetria):*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle \quad \forall\, \mathbf{v}, \mathbf{w} \in V\ ;$$

Um pouco de álgebra linear

Definição (Produto Interno)

Uma aplicação $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ *é um produto interno se satisfaz as seguintes propriedades:*

- *bilinearidade:*

$$\begin{aligned} \langle \mathbf{v}, \alpha\mathbf{w} + \beta\mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{z} \rangle \\ \langle \alpha\mathbf{v} + \beta\mathbf{w}, \mathbf{z} \rangle &= \alpha\langle \mathbf{v}, \mathbf{z} \rangle + \beta\langle \mathbf{w}, \mathbf{z} \rangle, \end{aligned}$$

  $\forall\, \alpha, \beta \in \mathbb{R}, \forall\, \mathbf{v}, \mathbf{w}, \mathbf{z} \in V;$

- *comutatividade (simetria):*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle \quad \forall\, \mathbf{v}, \mathbf{w} \in V\,;$$

- *positividade:*

$$\langle \mathbf{v}, \mathbf{v} \rangle > 0 \quad \forall\, \mathbf{v}, \quad e \quad \langle \mathbf{v}, \mathbf{v} \rangle = 0 \;\Leftrightarrow\; \mathbf{v} = \bar{0}$$

Um pouco de álgebra linear

Exemplo 7

No $\mathbb{R}^n$, o *produto interno usual* (produto escalar) dos vetores $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)^\top$ é definido por:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Um pouco de álgebra linear

Exemplo 7

No $\mathbb{R}^n$, o *produto interno usual* (produto escalar) dos vetores $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)^\top$ é definido por:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Exemplo 8

No espaço $M(n,n)$, um exemplo de produto interno entre as matrizes $\mathbf{A}, \mathbf{B} \in M(n,n)$ é dado por:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}$$

Exemplo 9

No espaço $\mathcal{C}([a,b])$ (espaço das funções contínuas no intervalo $[a,b]$), um exemplo de produto interno entre as funções contínuas $f,g:[a,b]\to\mathbb{R}$ é dado por:

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

Um pouco de álgebra linear

Definição (Norma)

Seja V um espaço vetorial. Uma aplicação $\|\cdot\| : V \longrightarrow \mathbb{R}$ que satisfaz as seguintes propriedades:

$$
\begin{aligned}
\|\mathbf{v}\| &\geq 0, \quad \forall\, \mathbf{v} \in V \\
\|\mathbf{v}\| &= 0 \Leftrightarrow \quad \mathbf{v} = \bar{0} \\
\|\lambda \mathbf{v}\| &= |\lambda|\|\mathbf{v}\|, \quad \forall\, \mathbf{v} \in V \quad e \quad \forall \lambda \in \mathbb{R}; \\
\|\mathbf{v} + \mathbf{w}\| &\leq \|\mathbf{v}\| + \|\mathbf{w}\|, \quad \forall\, \mathbf{v}, \mathbf{w} \in V
\end{aligned}
$$

é dita norma em V.

Um pouco de álgebra linear

Definição (Norma)

Seja V um espaço vetorial. Uma aplicação $\|\cdot\| : V \longrightarrow \mathbb{R}$ que satisfaz as seguintes propriedades:

$$\begin{aligned}
\|\mathbf{v}\| &\geq 0, \quad \forall\, \mathbf{v} \in V \\
\|\mathbf{v}\| &= 0 \;\Leftrightarrow\; \mathbf{v} = \bar{0} \\
\|\lambda \mathbf{v}\| &= |\lambda|\|\mathbf{v}\|, \quad \forall\, \mathbf{v} \in V \quad e \quad \forall \lambda \in \mathbb{R}; \\
\|\mathbf{v} + \mathbf{w}\| &\leq \|\mathbf{v}\| + \|\mathbf{w}\|, \quad \forall\, \mathbf{v}, \mathbf{w} \in V
\end{aligned}$$

é dita norma em V.

Definição (Norma Induzida pelo Produto Interno)

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Um pouco de álgebra linear

Exemplo 10

São exemplos de normas em $\mathbb{R}^n$:

1. $\|\mathbf{x}\|_\infty = \max\limits_{i=1\ldots n} \{|x_i|\}$ (norma do máximo)
2. $\|\mathbf{x}\|_1 = \sum\limits_{i=1}^{n} |x_i|$ (norma da soma)
3. $\|\mathbf{x}\|_2 = \left(\sum\limits_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$ (norma euclidiana)

Um pouco de álgebra linear

Exemplo 10

São exemplos de normas em $\mathbb{R}^n$:

1. $\|\mathbf{x}\|_\infty = \max\limits_{i=1\ldots n} \{|x_i|\}$ (norma do máximo)
2. $\|\mathbf{x}\|_1 = \sum\limits_{i=1}^{n} |x_i|$ (norma da soma)
3. $\|\mathbf{x}\|_2 = \left(\sum\limits_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$ (norma euclidiana)

Em particular, a norma $\|\cdot\|_2$ provém do produto interno

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

Um pouco de álgebra linear

Exemplo 11

Seja $\mathbf{x} = (1, 2, -5)^\top$

Um pouco de álgebra linear

Exemplo 11

Seja $\mathbf{x} = (1, 2, -5)^\top$

1. $\|\mathbf{x}\|_\infty = \max\{|1|, |2|, |-5|\} = 5$

Um pouco de álgebra linear

Exemplo 11

Seja $\mathbf{x} = (1, 2, -5)^\top$

1. $\|\mathbf{x}\|_\infty = \max\{|1|, |2|, |-5|\} = 5$
2. $\|\mathbf{x}\|_1 = |1| + |2| + |-5| = 8$

Um pouco de álgebra linear

Exemplo 11

Seja $\mathbf{x} = (1, 2, -5)^\top$

1. $\|\mathbf{x}\|_\infty = \max\{|1|, |2|, |-5|\} = 5$
2. $\|\mathbf{x}\|_1 = |1| + |2| + |-5| = 8$
3. $\|\mathbf{x}\|_2 = \sqrt{1^2 + 2^2 + (-5)^2} = \sqrt{30} \approx 5.48$

Um pouco de álgebra linear

Exemplo 11

Seja $\mathbf{x} = (1, 2, -5)^\top$

1. $\|\mathbf{x}\|_\infty = \max\{|1|, |2|, |-5|\} = 5$
2. $\|\mathbf{x}\|_1 = |1| + |2| + |-5| = 8$
3. $\|\mathbf{x}\|_2 = \sqrt{1^2 + 2^2 + (-5)^2} = \sqrt{30} \approx 5.48$

```
y = norm(x,inf);
y = norm(x,1);
y = norm(x,2);
```
(ou simplesmente `norm(x)`)

Um pouco de álgebra linear

Exemplo 12

São exemplos de normas em $M(m,n)$:

1. $\|\mathbf{A}\|_\infty = \max\limits_{i=1\dots m}\{\sum_{j=1}^{n}|a_{ij}|\}$ (norma linha)
2. $\|\mathbf{A}\|_1 = \max\limits_{j=1\dots n}\{\sum_{i=1}^{m}|a_{ij}|\}$ (norma coluna)
3. $\|\mathbf{A}\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}^2\right)^{\frac{1}{2}}$ (norma de Frobenius)

Um pouco de álgebra linear

Exemplo 12

São exemplos de normas em $M(m,n)$:

1. $\|\mathbf{A}\|_\infty = \max\limits_{i=1\dots m}\{\sum\limits_{j=1}^{n}|a_{ij}|\}$ (norma linha)
2. $\|\mathbf{A}\|_1 = \max\limits_{j=1\dots n}\{\sum\limits_{i=1}^{m}|a_{ij}|\}$ (norma coluna)
3. $\|\mathbf{A}\|_F = \left(\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}a_{ij}^2\right)^{\frac{1}{2}}$ (norma de Frobenius)

Em particular, a norma $\|\cdot\|_F$ provém do produto interno

$$\langle \mathbf{A}, \mathbf{B}\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, b_{ij}$$

Um pouco de álgebra linear

Exemplo 13

Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 4 & 0 & 2 \end{bmatrix}$

Um pouco de álgebra linear

Exemplo 13

Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 4 & 0 & 2 \end{bmatrix}$

1. $\|\mathbf{A}\|_\infty = \max\{4, 4, 6\} = 6$

Um pouco de álgebra linear

Exemplo 13

Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 4 & 0 & 2 \end{bmatrix}$

1. $\|\mathbf{A}\|_\infty = \max\{4, 4, 6\} = 6$
2. $\|\mathbf{A}\|_1 = \max\{7, 1, 6\} = 7$

Um pouco de álgebra linear

Exemplo 13

Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 4 & 0 & 2 \end{bmatrix}$

1. $||\mathbf{A}||_\infty = \max\{4, 4, 6\} = 6$
2. $||\mathbf{A}||_1 = \max\{7, 1, 6\} = 7$
3. $||\mathbf{A}||_F = \sqrt{1^2 + (-3)^2 + (-2)^2 + 1^2 + 1^2 + 4^2 + 2^2} = \sqrt{36} = 6$

Um pouco de álgebra linear

Exemplo 13

Seja $\mathbf{A} = \begin{bmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 4 & 0 & 2 \end{bmatrix}$

1. $||\mathbf{A}||_\infty = \max\{4, 4, 6\} = 6$
2. $||\mathbf{A}||_1 = \max\{7, 1, 6\} = 7$
3. $||\mathbf{A}||_F = \sqrt{1^2 + (-3)^2 + (-2)^2 + 1^2 + 1^2 + 4^2 + 2^2} = \sqrt{36} = 6$

```
y = norm(A,inf);
y = norm(A,1);
y = norm(A,'fro');
```

Um pouco de álgebra linear

Exemplo 14

São exemplos de normas em $\mathcal{C}([a,b])$:

1. $\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$
2. $\|f\|_1 = \int_a^b |f(x)|\, dx$
3. $\|f\|_2 = \left(\int_a^b |f(x)|^2 \right)^{\frac{1}{2}}$

Um pouco de álgebra linear

Exemplo 14

São exemplos de normas em $\mathcal{C}([a,b])$:

1. $\|f\|_\infty = \max\limits_{x\in[a,b]} |f(x)|$
2. $\|f\|_1 = \int_a^b |f(x)|\, dx$
3. $\|f\|_2 = \left(\int_a^b |f(x)|^2\right)^{\frac{1}{2}}$

Em particular, a norma $\|\cdot\|_2$ provém do produto interno

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

Um pouco de álgebra linear

Definição (Distância)

Uma aplicação $dist : V \times V \to \mathbb{R}$ *é chamada de distância se:*

$$\begin{aligned} dist(\mathbf{v},\mathbf{w}) &= dist(\mathbf{w},\mathbf{v}), \quad \forall\, \mathbf{v}\,,\mathbf{w} \in V \\ dist(\mathbf{v},\mathbf{w}) &\geq 0\,, \quad \forall\, \mathbf{v},\mathbf{w} \in V \\ dist(\mathbf{v},\mathbf{w}) = 0 &\Leftrightarrow \mathbf{v} = \mathbf{w} \\ dist(\mathbf{v},\mathbf{z}) &\leq dist(\mathbf{v},\mathbf{w}) + dist(\mathbf{w},\mathbf{z})\,, \quad \forall\, \mathbf{v},\mathbf{w},\mathbf{z} \in V \end{aligned}$$

Um pouco de álgebra linear

Definição (Distância)

Uma aplicação $dist : V \times V \to \mathbb{R}$ *é chamada de distância se:*

$$\begin{aligned} dist(\mathbf{v},\mathbf{w}) &= dist(\mathbf{w},\mathbf{v}), \quad \forall\, \mathbf{v}\,,\mathbf{w} \in V \\ dist(\mathbf{v},\mathbf{w}) &\geq 0\,, \quad \forall\, \mathbf{v},\mathbf{w} \in V \\ dist(\mathbf{v},\mathbf{w}) = 0 &\Leftrightarrow \mathbf{v} = \mathbf{w} \\ dist(\mathbf{v},\mathbf{z}) &\leq dist(\mathbf{v},\mathbf{w}) + dist(\mathbf{w},\mathbf{z})\,, \quad \forall\, \mathbf{v},\mathbf{w},\mathbf{z} \in V \end{aligned}$$

Teorema

Se $\|\cdot\|$ *é norma, então* $dist(\mathbf{v},\mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$ *é uma distância.*

Um pouco de álgebra linear

Exemplo 15

São exemplos de distâncias no $\mathbb{R}^n$:

1. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|$
2. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{\frac{1}{2}}$
3. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_\infty = \max\limits_{i=1\ldots n} |x_i - y_i|$

Um pouco de álgebra linear

Exemplo 15

São exemplos de distâncias no $\mathbb{R}^n$:

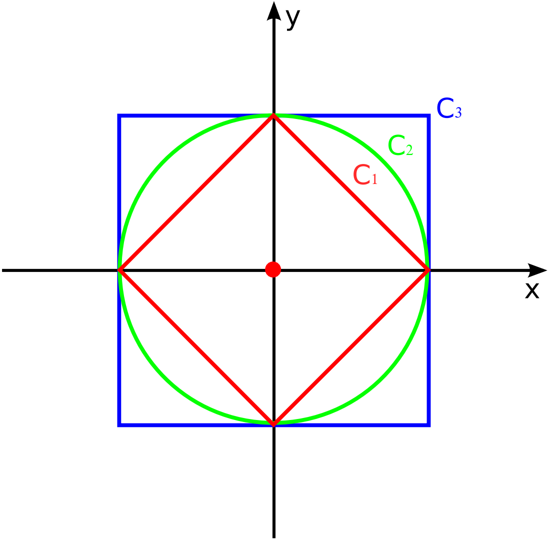
1. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|$
2. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{\frac{1}{2}}$
3. $dist(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_\infty = \max\limits_{i=1\ldots n} |x_i - y_i|$

Exemplo 15

Vejamos como ficaria o *disco unitário* $C = \{\mathbf{p} = (x, y) : dist(\mathbf{p}, \overline{0}) = 1\}$ em cada uma das distâncias acima:

1. $C_1 = \{(x, y) \in \mathbb{R}^2 : |x| + |y| = 1\}$
2. $C_2 = \{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} = 1\}$
3. $C_3 = \{(x, y) \in \mathbb{R}^2 : \max\{|x|, |y|\} = 1\}$

Um pouco de álgebra linear



Projeção ortogonal

Definição

Sejam $\mathbf{u}, \mathbf{v} \in E$ *vetores não nulos de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *e* $\mathbf{v}$ *são ditos ortogonais se* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. *Notação:* $\mathbf{u} \perp \mathbf{v}$.

Projeção ortogonal

Definição

Sejam $\mathbf{u}, \mathbf{v} \in E$ *vetores não nulos de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *e* $\mathbf{v}$ *são ditos ortogonais se* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$*. Notação:* $\mathbf{u} \perp \mathbf{v}$*.*

Exemplo 16

Os vetores do $\mathbb{R}^2$, $\mathbf{u} = (2,3)^\top$ e $\mathbf{v} = (-3,2)^\top$ são ortogonais. Pois,

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2 \times -3 + 3 \times 2 = 0.$$

Projeção ortogonal

Definição

Sejam $\mathbf{u}, \mathbf{v} \in E$ *vetores não nulos de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *e* $\mathbf{v}$ *são ditos ortogonais se* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$*. Notação:* $\mathbf{u} \perp \mathbf{v}$.

Exemplo 16

Os vetores do $\mathbb{R}^2$, $\mathbf{u} = (2,3)^\top$ e $\mathbf{v} = (-3,2)^\top$ são ortogonais. Pois,

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2 \times -3 + 3 \times 2 = 0.$$

Exemplo 17

As funções $\operatorname{sen}(x), \cos(x) \in \mathcal{C}([0, 2\pi])$ são ortogonais. Pois,

$$\langle \operatorname{sen}(x), \cos(x) \rangle = \int_0^{2\pi} \operatorname{sen}(x)\cos(x)dx = \frac{1}{2}\int_0^{2\pi} \operatorname{sen}(2x)dx = 0.$$

Projeção ortogonal

Definição

Seja $\mathbf{u} \in E$ *um vetor não nulo de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *é dito ortogonal a um subespaço* $V \subset E$*, se* $\mathbf{u} \perp \mathbf{v}, \forall \mathbf{v} \in V$*.*
Notação: $\mathbf{u} \perp V$*.*

Projeção ortogonal

Definição

Seja $\mathbf{u} \in E$ *um vetor não nulo de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *é dito ortogonal a um subespaço* $V \subset E$*, se* $\mathbf{u} \perp \mathbf{v}, \forall \mathbf{v} \in V$*. Notação:* $\mathbf{u} \perp V$*.*

Teorema

Se V *é um espaço de dimensão finita, e* $\mathcal{B} = \{\varphi_1, \dots, \varphi_n\}$ *uma base de* V*, então* $\mathbf{u} \perp V$ *se e somente se* $\mathbf{u} \perp \varphi_i \ \forall \ i = 1, \dots, n$*.*

Projeção ortogonal

Definição

Seja $\mathbf{u} \in E$ *um vetor não nulo de um espaço vetorial munido de produto interno. Então* $\mathbf{u}$ *é dito ortogonal a um subespaço* $V \subset E$*, se* $\mathbf{u} \perp \mathbf{v}, \forall \mathbf{v} \in V$*. Notação:* $\mathbf{u} \perp V$*.*

Teorema

Se V *é um espaço de dimensão finita, e* $\mathcal{B} = \{\varphi_1, \ldots, \varphi_n\}$ *uma base de* V*, então* $\mathbf{u} \perp V$ *se e somente se* $\mathbf{u} \perp \varphi_i \; \forall \, i = 1, \ldots, n$*.*

Definição (Projeção ortogonal)

Seja E *um espaço vetorial, e* $V \subset E$ *um subspaço de dimensão finita de* E*. A projeção ortogonal de* $\mathbf{u} \in E$ *sobre* V *é o vetor* $\mathbf{v}^* \in V$ *tal que* $\mathbf{u} - \mathbf{v}^* \perp V$*.*

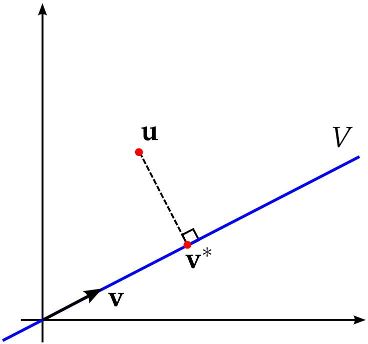
Melhor aproximação em $\mathbb{R}^2$

Problema

Sejam V um subespaço de dimensão 1 do $\mathbb{R}^2$ (uma reta que passa pela origem), e $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ um ponto qualquer fora de V. Encontrar um ponto $\mathbf{v}^* \in V$ que esteja mais próximo de $\mathbf{u}$.

Melhor aproximação em $\mathbb{R}^2$

Geometricamente:



Seja $V = \{\mathbf{r} = \lambda \mathbf{v}, \lambda \in \mathbb{R}\}$ a reta que passa pela origem, cujo vetor diretor é $\mathbf{v}$. Vamos encontrar o ponto $\mathbf{v}^* \in V$ tal que $\mathbf{u} - \mathbf{v}^*$ seja ortogonal à V.

Melhor aproximação em $\mathbb{R}^2$

Note que V é gerado pelo vetor $\mathbf{v}$, logo $\mathbf{v}^* = \alpha\mathbf{v}$. Portanto, basta encontrar $\alpha \in \mathbb{R}$ tal que:

$$\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v} \rangle = \langle \mathbf{u} - \alpha\mathbf{v}, \mathbf{v} \rangle = 0$$

Melhor aproximação em $\mathbb{R}^2$

Note que V é gerado pelo vetor $\mathbf{v}$, logo $\mathbf{v}^* = \alpha\mathbf{v}$. Portanto, basta encontrar $\alpha \in \mathbb{R}$ tal que:

$$\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v} \rangle = \langle \mathbf{u} - \alpha\mathbf{v}, \mathbf{v} \rangle = 0$$

Sendo $\langle \cdot, \cdot \rangle$ o produto usual em $\mathbb{R}^2$, segue que:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Melhor aproximação em $\mathbb{R}^2$

Note que V é gerado pelo vetor $\mathbf{v}$, logo $\mathbf{v}^* = \alpha\mathbf{v}$. Portanto, basta encontrar $\alpha \in \mathbb{R}$ tal que:

$$\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v} \rangle = \langle \mathbf{u} - \alpha\mathbf{v}, \mathbf{v} \rangle = 0$$

Sendo $\langle \cdot, \cdot \rangle$ o produto usual em $\mathbb{R}^2$, segue que:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$$

resultando

$$\mathbf{v}^* = \mathrm{pr}_{\mathbf{v}}(\mathbf{u}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}$$

Melhor aproximação em espaços vetoriais

Teorema

Seja V um espaço de dimensão finita de um espaço vetorial E. Se $\mathbf{u} \in E$, então $\mathbf{v}^* \in V$, a projeção ortogonal de $\mathbf{u}$ em V, é a melhor aproximação de $\mathbf{u}$ em V, no sentido que

$$\|\mathbf{u} - \mathbf{v}^*\| < \|\mathbf{u} - \mathbf{v}\|,$$

para qualquer que seja $\mathbf{v} \in V$ com $\mathbf{v} \neq \mathbf{v}^*$. Este pode ser escrito como

$$\mathbf{v}^* = \alpha_0^*\varphi_0 + \alpha_1^*\varphi_1 + \cdots + \alpha_n^*\varphi_n$$

onde o vetor $\alpha^* = (\alpha_0^*, \alpha_1^*, \ldots, \alpha_n^*)^\top$ é a solução do sistema linear:

$$\begin{bmatrix} \langle\varphi_0,\varphi_0\rangle & \cdots & \langle\varphi_0,\varphi_n\rangle \\ \vdots & & \vdots \\ \langle\varphi_n,\varphi_0\rangle & \cdots & \langle\varphi_n,\varphi_n\rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \vdots \\ \alpha_n^* \end{bmatrix} = \begin{bmatrix} \langle\mathbf{u},\varphi_0\rangle \\ \vdots \\ \langle\mathbf{u},\varphi_n\rangle \end{bmatrix}$$

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u} - \mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v} - \mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u} - \mathbf{v}^*$.

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u} - \mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v} - \mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u} - \mathbf{v}^*$. Daí,

$$\|\mathbf{u} - \mathbf{v}\|^2 = \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^*, \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^* \rangle$$

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u}-\mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v}-\mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u}-\mathbf{v}^*$. Daí,

$$\begin{aligned}
\|\mathbf{u}-\mathbf{v}\|^2 &= \langle \mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle = \langle \mathbf{u}-\mathbf{v}+\mathbf{v}^*-\mathbf{v}^*, \mathbf{u}-\mathbf{v}+\mathbf{v}^*-\mathbf{v}^*\rangle \\
&= \langle \mathbf{u}-\mathbf{v}^*, \mathbf{u}-\mathbf{v}^*\rangle + 2\langle \mathbf{u}-\mathbf{v}^*, \mathbf{v}^*-\mathbf{v}\rangle + \langle \mathbf{v}^*-\mathbf{v}, \mathbf{v}^*-\mathbf{v}\rangle
\end{aligned}$$

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u} - \mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v} - \mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u} - \mathbf{v}^*$. Daí,

$$\begin{aligned}
\|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^*, \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^* \rangle \\
&= \langle \mathbf{u} - \mathbf{v}^*, \mathbf{u} - \mathbf{v}^* \rangle + 2\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v}^* - \mathbf{v} \rangle + \langle \mathbf{v}^* - \mathbf{v}, \mathbf{v}^* - \mathbf{v} \rangle \\
&= \|\mathbf{u} - \mathbf{v}^*\|^2 + \|\mathbf{v}^* - \mathbf{v}\|^2 > \|\mathbf{u} - \mathbf{v}^*\|^2
\end{aligned}$$

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u} - \mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v} - \mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u} - \mathbf{v}^*$. Daí,

$$
\begin{aligned}
\|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^*, \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^* \rangle \\
&= \langle \mathbf{u} - \mathbf{v}^*, \mathbf{u} - \mathbf{v}^* \rangle + 2\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v}^* - \mathbf{v} \rangle + \langle \mathbf{v}^* - \mathbf{v}, \mathbf{v}^* - \mathbf{v} \rangle \\
&= \|\mathbf{u} - \mathbf{v}^*\|^2 + \|\mathbf{v}^* - \mathbf{v}\|^2 > \|\mathbf{u} - \mathbf{v}^*\|^2
\end{aligned}
$$

Portanto podemos concluir que $\|\mathbf{u} - \mathbf{v}^*\| < \|\mathbf{u} - \mathbf{v}\|, \forall\, \mathbf{v} \in V$. Ainda falta mostrar que α^* é solução do sistema linear $\mathbf{A}\alpha^* = \mathbf{b}$.

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u}-\mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v}-\mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u}-\mathbf{v}^*$. Daí,

$$\begin{aligned}
\|\mathbf{u}-\mathbf{v}\|^2 &= \langle \mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle = \langle \mathbf{u}-\mathbf{v}+\mathbf{v}^*-\mathbf{v}^*, \mathbf{u}-\mathbf{v}+\mathbf{v}^*-\mathbf{v}^*\rangle \\
&= \langle \mathbf{u}-\mathbf{v}^*, \mathbf{u}-\mathbf{v}^*\rangle + 2\langle \mathbf{u}-\mathbf{v}^*, \mathbf{v}^*-\mathbf{v}\rangle + \langle \mathbf{v}^*-\mathbf{v}, \mathbf{v}^*-\mathbf{v}\rangle \\
&= \|\mathbf{u}-\mathbf{v}^*\|^2 + \|\mathbf{v}^*-\mathbf{v}\|^2 > \|\mathbf{u}-\mathbf{v}^*\|^2
\end{aligned}$$

Portanto podemos concluir que $\|\mathbf{u}-\mathbf{v}^*\| < \|\mathbf{u}-\mathbf{v}\|, \forall\, \mathbf{v} \in V$. Ainda falta mostrar que α^* é solução do sistema linear $\mathbf{A}\alpha^* = \mathbf{b}$.

Seja $\mathcal{B} = \{\varphi_0, \ldots, \varphi_n\}$ base de V. Como $\mathbf{u}-\mathbf{v}^* \perp V$, temos que $\langle \mathbf{u}-\mathbf{v}^*, \varphi_i\rangle = 0,\ \forall i = 0, \ldots, n$. Assim, como $\mathbf{v}^* \in V$,

$$\langle \mathbf{u} - (\alpha_0^*\varphi_0 + \cdots + \alpha_n^*\varphi_n), \varphi_i\rangle \quad = \quad 0$$

Melhor aproximação em espaços vetoriais

Demonstração: Seja $\mathbf{v}^*$ a projeção ortogonal de $\mathbf{u}$ em V, ou seja, $\mathbf{u} - \mathbf{v}^* \perp V$. Seja $\mathbf{v} \in V$ um vetor de V, $\mathbf{v} \neq \mathbf{v}^*$. Então $\mathbf{v} - \mathbf{v}^* \in V$, e portanto, ortogonal a $\mathbf{u} - \mathbf{v}^*$. Daí,

$$\begin{aligned}
\|\mathbf{u} - \mathbf{v}\|^2 &= \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \langle \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^*, \mathbf{u} - \mathbf{v} + \mathbf{v}^* - \mathbf{v}^* \rangle \\
&= \langle \mathbf{u} - \mathbf{v}^*, \mathbf{u} - \mathbf{v}^* \rangle + 2\langle \mathbf{u} - \mathbf{v}^*, \mathbf{v}^* - \mathbf{v} \rangle + \langle \mathbf{v}^* - \mathbf{v}, \mathbf{v}^* - \mathbf{v} \rangle \\
&= \|\mathbf{u} - \mathbf{v}^*\|^2 + \|\mathbf{v}^* - \mathbf{v}\|^2 > \|\mathbf{u} - \mathbf{v}^*\|^2
\end{aligned}$$

Portanto podemos concluir que $\|\mathbf{u} - \mathbf{v}^*\| < \|\mathbf{u} - \mathbf{v}\|, \forall\, \mathbf{v} \in V$. Ainda falta mostrar que α^* é solução do sistema linear $\mathbf{A}\alpha^* = \mathbf{b}$.

Seja $\mathcal{B} = \{\varphi_0, \ldots, \varphi_n\}$ base de V. Como $\mathbf{u} - \mathbf{v}^* \perp V$, temos que $\langle \mathbf{u} - \mathbf{v}^*, \varphi_i \rangle = 0,\ \forall i = 0, \ldots, n$. Assim, como $\mathbf{v}^* \in V$,

$$\begin{aligned}
\langle \mathbf{u} - (\alpha_0^*\varphi_0 + \cdots + \alpha_n^*\varphi_n), \varphi_i \rangle &= 0 \quad \Leftrightarrow \\
\langle \varphi_i, \varphi_0 \rangle \alpha_0^* + \cdots + \langle \varphi_i, \varphi_n \rangle \alpha_n^* &= \langle \mathbf{u}, \varphi_i \rangle, \quad \forall i = 0, \ldots, n
\end{aligned}$$

o que conclui a demonstração. ■

Melhor aproximação em espaços vetoriais

Proposição

O sistema linear normal:

$$\begin{bmatrix} \langle \varphi_0, \varphi_0 \rangle & \cdots & \langle \varphi_0, \varphi_n \rangle \\ \vdots & & \vdots \\ \langle \varphi_n, \varphi_0 \rangle & \cdots & \langle \varphi_n, \varphi_n \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \vdots \\ \alpha_n^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}, \varphi_0 \rangle \\ \vdots \\ \langle \mathbf{u}, \varphi_n \rangle \end{bmatrix}$$

possui uma única solução.

Melhor aproximação em espaços vetoriais

Demonstração:

Melhor aproximação em espaços vetoriais

Demonstração:

Basta mostrar que $\det(\mathbf{A}) \neq 0$, com $a_{ij} = \langle \varphi_i, \varphi_j \rangle$.

Melhor aproximação em espaços vetoriais

Demonstração:
Basta mostrar que $\det(\mathbf{A}) \neq 0$, com $a_{ij} = \langle \varphi_i, \varphi_j \rangle$.

Considere o sistema linear homogêneo $\mathbf{A}\beta = \overline{0}$. Logo,

Melhor aproximação em espaços vetoriais

Demonstração:
Basta mostrar que $\det(\mathbf{A}) \neq 0$, com $a_{ij} = \langle \varphi_i, \varphi_j \rangle$.

Considere o sistema linear homogêneo $\mathbf{A}\beta = \overline{0}$. Logo,

$$\begin{aligned} 0 &= \beta \cdot \mathbf{A}\beta = \textstyle\sum_{i,j=0}^{n} \beta_i a_{ij} \beta_j = \sum_{i,j=0}^{n} \beta_i \left\langle \varphi_i, \varphi_j \right\rangle \beta_j \\ &= \left\langle \textstyle\sum_{i=0}^{n} \beta_i \varphi_i, \sum_{j=0}^{n} \beta_j \varphi_j \right\rangle = \left\| \textstyle\sum_{i=0}^{n} \beta_i \varphi_i \right\|^2 \end{aligned}$$

Melhor aproximação em espaços vetoriais

Demonstração:
Basta mostrar que $\det(\mathbf{A}) \neq 0$, com $a_{ij} = \langle \varphi_i, \varphi_j \rangle$.

Considere o sistema linear homogêneo $\mathbf{A}\beta = \bar{0}$. Logo,

$$\begin{aligned} 0 &= \beta \cdot \mathbf{A}\beta = \textstyle\sum_{i,j=0}^{n} \beta_i a_{ij} \beta_j = \sum_{i,j=0}^{n} \beta_i \left\langle \varphi_i, \varphi_j \right\rangle \beta_j \\ &= \left\langle \textstyle\sum_{i=0}^{n} \beta_i \varphi_i, \sum_{j=0}^{n} \beta_j \varphi_j \right\rangle = \left\| \textstyle\sum_{i=0}^{n} \beta_i \varphi_i \right\|^2 \end{aligned}$$

Pelo fato de $\sum_{i=0}^{n} \beta_i \varphi_i = \bar{0}$ e $\{\varphi_0, \ldots, \varphi_n\}$ ser uma base, temos que $\beta_0 = \cdots = \beta_n = 0 \Rightarrow \beta = \bar{0}$. Logo, o sistema linear homogêneo possui apenas a solução trivial.

Melhor aproximação em espaços vetoriais

Demonstração:
Basta mostrar que $\det(\mathbf{A}) \neq 0$, com $a_{ij} = \langle \varphi_i, \varphi_j \rangle$.

Considere o sistema linear homogêneo $\mathbf{A}\beta = \bar{0}$. Logo,

$$\begin{aligned} 0 &= \beta \cdot \mathbf{A}\beta = \textstyle\sum_{i,j=0}^{n} \beta_i a_{ij} \beta_j = \sum_{i,j=0}^{n} \beta_i \left\langle \varphi_i, \varphi_j \right\rangle \beta_j \\ &= \left\langle \textstyle\sum_{i=0}^{n} \beta_i \varphi_i, \sum_{j=0}^{n} \beta_j \varphi_j \right\rangle = \left\| \textstyle\sum_{i=0}^{n} \beta_i \varphi_i \right\|^2 \end{aligned}$$

Pelo fato de $\sum_{i=0}^{n} \beta_i \varphi_i = \bar{0}$ e $\{\varphi_0, \ldots, \varphi_n\}$ ser uma base, temos que $\beta_0 = \cdots = \beta_n = 0 \Rightarrow \beta = \bar{0}$. Logo, o sistema linear homogêneo possui apenas a solução trivial.

Portanto, $\det(\mathbf{A}) \neq 0$. ∎

Desigualdade de Cauchy-Schwarz

Seja V um espaço vetorial munido de produto interno, então:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \forall \mathbf{u}, \mathbf{v} \in V$$

Desigualdade de Cauchy-Schwarz

Seja V um espaço vetorial munido de produto interno, então:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \ \forall \mathbf{u}, \mathbf{v} \in V$$

Demonstração:

Caso $\mathbf{u} = \overline{0}$ ou $\mathbf{v} = \overline{0}$, temos que $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$.

Desigualdade de Cauchy-Schwarz

Seja V um espaço vetorial munido de produto interno, então:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \forall \mathbf{u}, \mathbf{v} \in V$$

Demonstração:

Caso $\mathbf{u} = \bar{0}$ ou $\mathbf{v} = \bar{0}$, temos que $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$.

Se $\mathbf{u} \neq \bar{0}$ e $\mathbf{v} \neq \bar{0}$, fazendo a projeção ortogonal de $\mathbf{u}$ em $\mathbf{v}$ temos:

Desigualdade de Cauchy-Schwarz

Seja V um espaço vetorial munido de produto interno, então:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \ \forall \mathbf{u}, \mathbf{v} \in V$$

Demonstração:

Caso $\mathbf{u} = \bar{0}$ ou $\mathbf{v} = \bar{0}$, temos que $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$.

Se $\mathbf{u} \neq \bar{0}$ e $\mathbf{v} \neq \bar{0}$, fazendo a projeção ortogonal de $\mathbf{u}$ em $\mathbf{v}$ temos:

$$\|\mathrm{pr}_{\mathbf{v}}(\mathbf{u})\| \leq \|\mathbf{u}\| \Leftrightarrow \left\| \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \mathbf{v} \right\| \leq \|\mathbf{u}\| \Leftrightarrow \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{\|\mathbf{v}\|^2} \|\mathbf{v}\| \leq \|\mathbf{u}\|$$

Desigualdade de Cauchy-Schwarz

Seja V um espaço vetorial munido de produto interno, então:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \; \forall \mathbf{u}, \mathbf{v} \in V$$

Demonstração:

Caso $\mathbf{u} = \overline{0}$ ou $\mathbf{v} = \overline{0}$, temos que $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$.

Se $\mathbf{u} \neq \overline{0}$ e $\mathbf{v} \neq \overline{0}$, fazendo a projeção ortogonal de $\mathbf{u}$ em $\mathbf{v}$ temos:

$$\|\mathrm{pr}_{\mathbf{v}}(\mathbf{u})\| \leq \|\mathbf{u}\| \Leftrightarrow \left\| \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \mathbf{v} \right\| \leq \|\mathbf{u}\| \Leftrightarrow \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{\|\mathbf{v}\|^2} \|\mathbf{v}\| \leq \|\mathbf{u}\|$$

Portanto, $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|, \; \forall \mathbf{u}, \mathbf{v} \in V$.

■

O método dos mínimos quadrados

Seja f uma função, e V um espaço de funções conhecidas, de dimensão $n+1<\infty$, gerado pelas funções $\{\varphi_0,\ldots,\varphi_n\}$. Desejamos encontrar a função $F^*\in V$, $F^*=\alpha_0^*\varphi_0+\cdots+\alpha_n^*\varphi_n$, que melhor aproxima a função f, isto é, queremos encontrar F^* que minimize:

$$dist(f,F^*)$$

O método dos mínimos quadrados

Seja f uma função, e V um espaço de funções conhecidas, de dimensão $n+1<\infty$, gerado pelas funções $\{\varphi_0,\dots,\varphi_n\}$. Desejamos encontrar a função $F^*\in V$, $F^*=\alpha_0^*\varphi_0+\cdots+\alpha_n^*\varphi_n$, que melhor aproxima a função f, isto é, queremos encontrar F^* que minimize:

$$dist(f,F^*)$$

Na verdade, queremos obter:

$$Q=\min_{\alpha^*}\|f-F^*\|^2$$

Daí, o nome de *mínimos quadrados*.

O método dos mínimos quadrados

Portanto, para encontrar a função:

$$F^* = \alpha_0^* \varphi_0 + \cdots + \alpha_n^* \varphi_n$$

que melhor aproxima a função f no sentido dos mínimos quadrados, basta calcular e resolver o sistema de equações normais:

$$\begin{bmatrix} \langle \varphi_0, \varphi_0 \rangle & \cdots & \langle \varphi_0, \varphi_n \rangle \\ \vdots & & \vdots \\ \langle \varphi_n, \varphi_0 \rangle & \cdots & \langle \varphi_n, \varphi_n \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \vdots \\ \alpha_n^* \end{bmatrix} = \begin{bmatrix} \langle f, \varphi_0 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{bmatrix}.$$

O método dos mínimos quadrados

Exercício 1

Deduza o sistema de equações normais via otimização. Use o fato de que α^* é ponto crítico de Q, isto é, $\nabla Q = \bar{0}$.

Aproximação polinomial: caso contínuo

Dada uma função $f \in \mathcal{C}([a,b])$. Desejamos aproximar $f(x)$ por um polinômio $P_m \in V = \mathcal{P}_m$, isto é:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x)$$

de tal forma que minimize:

Aproximação polinomial: caso contínuo

Dada uma função $f \in \mathcal{C}([a,b])$. Desejamos aproximar $f(x)$ por um polinômio $P_m \in V = \mathcal{P}_m$, isto é:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x)$$

de tal forma que minimize:

$$Q = \|f - P_m\|^2 = \langle f - P_m, f - P_m \rangle = \int_a^b (f(x) - P_m(x))^2 \, dx$$

Aproximação polinomial: caso contínuo

Dada uma função $f \in \mathcal{C}([a,b])$. Desejamos aproximar $f(x)$ por um polinômio $P_m \in V = \mathcal{P}_m$, isto é:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x)$$

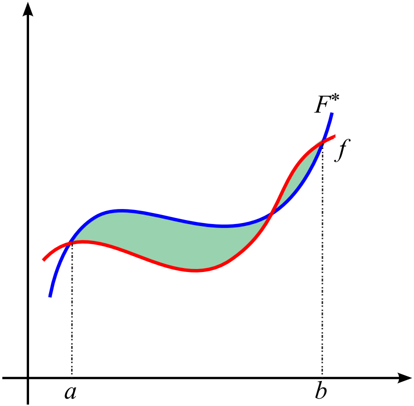
de tal forma que minimize:

$$Q = \|f - P_m\|^2 = \langle f - P_m, f - P_m \rangle = \int_a^b (f(x) - P_m(x))^2 \, dx$$

Por outro lado, sabemos que $\mathcal{B} = \{1, x, x^2, \ldots, x^m\}$ é uma base de $\mathcal{P}_m$. Assim, para obter $P_m(x)$ basta resolver o sistema linear:

$$\begin{bmatrix} \langle 1,1 \rangle & \langle 1,x \rangle & \cdots & \langle 1,x^m \rangle \\ \langle x,1 \rangle & \langle x,x \rangle & \cdots & \langle x,x^m \rangle \\ \vdots & \vdots & \cdots & \vdots \\ \langle x^m,1 \rangle & \langle x^m,x \rangle & \cdots & \langle x^m,x^m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle 1,f \rangle \\ \langle x,f \rangle \\ \vdots \\ \langle x^m,f \rangle \end{bmatrix}$$

Interpretação geométrica: caso contínuo



Minimizar o quadrado da área formada entre as funções.

Aproximação polinomial: caso contínuo

Exemplo 17

Seja $f(x) = e^x$, com $x \in [0,1]$. Queremos aproximar $f(x)$ por uma reta, ou seja, $f(x) \approx \alpha_0^* + \alpha_1^* x = P_1(x)$. Vamos considerar o produto escalar

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\, dx$$

Assim, para determinar $P_1 \in \mathcal{P}_1$ que melhor aproxima f no sentido de mínimos quadrados, primeiramente é preciso calcular a matriz $\mathbf{A}$ e o vetor $\mathbf{b}$, e depois resolver $\mathbf{A}\alpha^* = \mathbf{b}$.

Aproximação polinomial: caso contínuo

Exemplo 17

Seja $f(x) = e^x$, com $x \in [0,1]$. Queremos aproximar $f(x)$ por uma reta, ou seja, $f(x) \approx \alpha_0^* + \alpha_1^* x = P_1(x)$. Vamos considerar o produto escalar

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\, dx$$

Assim, para determinar $P_1 \in \mathcal{P}_1$ que melhor aproxima f no sentido de mínimos quadrados, primeiramente é preciso calcular a matriz $\mathbf{A}$ e o vetor $\mathbf{b}$, e depois resolver $\mathbf{A}\alpha^* = \mathbf{b}$. Logo,

$$\begin{array}{ll}
\langle 1,1 \rangle = \int_0^1 1 \cdot 1\, dx = 1 & \langle 1,x \rangle = \int_0^1 1 \cdot x\, dx = \frac{1}{2} \\
\langle x,x \rangle = \int_0^1 x \cdot x\, dx = \frac{1}{3} & \langle x,1 \rangle = \langle 1,x \rangle \\
\langle 1,f \rangle = \int_0^1 1 \cdot e^x\, dx = e - 1 & \langle x,f \rangle = \int_0^1 x \cdot e^x\, dx = 1
\end{array}$$

Aproximação polinomial: caso contínuo

Continuação

chegando-se ao sistema linear

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \end{bmatrix} = \begin{bmatrix} e-1 \\ 1 \end{bmatrix}$$

cuja solução é $\alpha_0^* = 0.873$ e $\alpha_1^* = 1.690$. Assim, a aproximação procurada é

$$P_1(x) = \alpha_0^* + \alpha_1^* x = 0.873 + 1.690x$$

Erro de Truncamento: caso contínuo

O *erro truncamento* da aproximação f por F^* no sentido de mínimos quadrados é dado por $Q = \|f - F^*\|^2$, isto é, o quadrado da distância de f à sua aproximação F^*.

Erro de Truncamento: caso contínuo

O *erro truncamento* da aproximação f por F^* no sentido de mínimos quadrados é dado por $Q = \|f - F^*\|^2$, isto é, o quadrado da distância de f à sua aproximação F^*.

Exemplo 18

Vamos calcular o erro de truncamento do exemplo anterior. Assim,

$$Q = \|f - P_1\|^2 \quad = \quad \langle f - P_1, f - P_1 \rangle$$

Erro de Truncamento: caso contínuo

O *erro truncamento* da aproximação f por F^* no sentido de mínimos quadrados é dado por $Q = \|f - F^*\|^2$, isto é, o quadrado da distância de f à sua aproximação F^*.

Exemplo 18

Vamos calcular o erro de truncamento do exemplo anterior. Assim,

$$\begin{aligned} Q = \|f - P_1\|^2 &= \langle f - P_1, f - P_1 \rangle \\ &= \langle f - \alpha_0^* - \alpha_1^* x, f - \alpha_0^* - \alpha_1^* x \rangle \end{aligned}$$

Erro de Truncamento: caso contínuo

O *erro truncamento* da aproximação f por F^* no sentido de mínimos quadrados é dado por $Q = \|f - F^*\|^2$, isto é, o quadrado da distância de f à sua aproximação F^*.

Exemplo 18

Vamos calcular o erro de truncamento do exemplo anterior. Assim,

$$\begin{aligned} Q = \|f - P_1\|^2 &= \langle f - P_1, f - P_1 \rangle \\ &= \langle f - \alpha_0^* - \alpha_1^* x, f - \alpha_0^* - \alpha_1^* x \rangle \\ &= \langle f, f \rangle - 2\alpha_0^* \langle f, 1 \rangle - 2\alpha_1^* \langle f, x \rangle + (\alpha_0^*)^2 \langle 1, 1 \rangle \\ &+ 2\alpha_0^* \alpha_1^* \langle 1, x \rangle + (\alpha_1^*)^2 \langle x, x \rangle \end{aligned}$$

Erro de Truncamento: caso contínuo

O *erro truncamento* da aproximação f por F^* no sentido de mínimos quadrados é dado por $Q = \|f - F^*\|^2$, isto é, o quadrado da distância de f à sua aproximação F^*.

Exemplo 18

Vamos calcular o erro de truncamento do exemplo anterior. Assim,

$$\begin{aligned}
Q = \|f - P_1\|^2 &= \langle f - P_1, f - P_1 \rangle \\
&= \langle f - \alpha_0^* - \alpha_1^* x, f - \alpha_0^* - \alpha_1^* x \rangle \\
&= \langle f, f \rangle - 2\alpha_0^* \langle f, 1 \rangle - 2\alpha_1^* \langle f, x \rangle + (\alpha_0^*)^2 \langle 1, 1 \rangle \\
&+ 2\alpha_0^* \alpha_1^* \langle 1, x \rangle + (\alpha_1^*)^2 \langle x, x \rangle \\
&= 0.003969
\end{aligned}$$

Aproximação polinomial: caso discreto

Dada uma função $f(x)$ amostrada , ou seja, é conhecida apenas nos $(n+1)$ pares de pontos:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

onde $y_i = f(x_i)$, $i = 0, \ldots, n$, com os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Aproximação polinomial: caso discreto

Dada uma função $f(x)$ amostrada , ou seja, é conhecida apenas nos $(n+1)$ pares de pontos:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

onde $y_i = f(x_i)$, $i = 0, \ldots, n$, com os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Desejamos aproximar a função f por um polinômio $P_m \in \mathcal{P}_m$, isto é:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x)$$

com $m < n$, tal que:

$$Q = \min_{\alpha^*} \|f - P_m\|^2$$

Aproximação polinomial: caso discreto

Dada uma função $f(x)$ amostrada , ou seja, é conhecida apenas nos $(n+1)$ pares de pontos:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

onde $y_i = f(x_i)$, $i = 0, \ldots, n$, com os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Desejamos aproximar a função f por um polinômio $P_m \in \mathcal{P}_m$, isto é:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x)$$

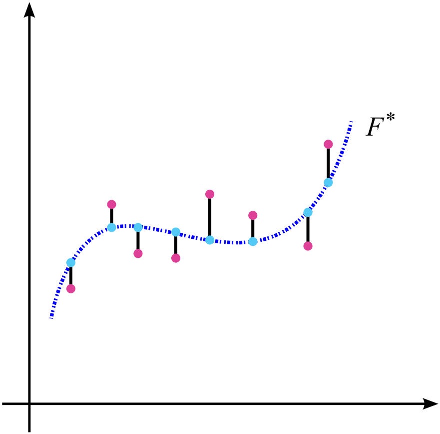
com $m < n$, tal que:

$$Q = \min_{\alpha^*} \|f - P_m\|^2$$

Para isso vamos utilizar o seguinte produto interno usal do $\mathbb{R}^{n+1}$ (produto escalar) :

$$\langle f, g \rangle = \sum_{k=0}^{n} f(x_k) g(x_k)$$

Interpretação geométrica: caso discreto



Minimizar a soma dos desvios $|y_k - F^*(x_k)|$ ao quadrado.

Aproximação polinomial: caso discreto

Utilizando o produto interno usal do $\mathbb{R}^{n+1}$, temos:

$$Q = \|f - P_m\|^2 = \langle f - P_m, f - P_m \rangle = \sum_{k=0}^{n} (y_k - P_m(x_k))^2$$

Aproximação polinomial: caso discreto

Utilizando o produto interno usal do $\mathbb{R}^{n+1}$, temos:

$$Q = \|f - P_m\|^2 = \langle f - P_m, f - P_m \rangle = \sum_{k=0}^{n} \left(y_k - P_m(x_k)\right)^2$$

Agora, vamos supor que $\mathbf{y} = (y_0, y_1, \ldots, y_n)^\top \in \mathbb{R}^{n+1}$ e

$$\mathbf{p} = \left(P_m(x_0), P_m(x_1), \ldots, P_m(x_n)\right)^\top \in \mathbb{R}^{n+1}$$

Aproximação polinomial: caso discreto

Utilizando o produto interno usal do $\mathbb{R}^{n+1}$, temos:

$$Q = \|f - P_m\|^2 = \langle f - P_m, f - P_m \rangle = \sum_{k=0}^{n} (y_k - P_m(x_k))^2$$

Agora, vamos supor que $\mathbf{y} = (y_0, y_1, \ldots, y_n)^\top \in \mathbb{R}^{n+1}$ e

$$\mathbf{p} = (P_m(x_0), P_m(x_1), \ldots, P_m(x_n))^\top \in \mathbb{R}^{n+1}$$

Por outro lado,

$$\begin{aligned}
P_m(x_0) &= \alpha_0^* + \alpha_1^* x_0 + \alpha_2^* x_0^2 + \cdots + \alpha_m^* x_0^m \\
P_m(x_1) &= \alpha_0^* + \alpha_1^* x_1 + \alpha_2^* x_1^2 + \cdots + \alpha_m^* x_1^m \\
&\vdots \\
P_m(x_n) &= \alpha_0^* + \alpha_1^* x_n + \alpha_2^* x_n^2 + \cdots + \alpha_m^* x_n^m
\end{aligned}$$

Aproximação polinomial: caso discreto

Assim, podemos reescrever o vetor **p** da seguinte forma:

Aproximação polinomial: caso discreto

Assim, podemos reescrever o vetor **p** da seguinte forma:

$$\mathbf{p} = \alpha_0^* \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{u}_0} + \alpha_1^* \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{u}_1} + \alpha_2^* \underbrace{\begin{bmatrix} x_0^2 \\ x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}}_{\mathbf{u}_2} + \cdots + \alpha_m^* \underbrace{\begin{bmatrix} x_0^m \\ x_1^m \\ \vdots \\ x_n^m \end{bmatrix}}_{\mathbf{u}_m}$$

Aproximação polinomial: caso discreto

Assim, podemos reescrever o vetor **p** da seguinte forma:

$$\mathbf{p} = \alpha_0^* \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{u}_0} + \alpha_1^* \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{u}_1} + \alpha_2^* \underbrace{\begin{bmatrix} x_0^2 \\ x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}}_{\mathbf{u}_2} + \cdots + \alpha_m^* \underbrace{\begin{bmatrix} x_0^m \\ x_1^m \\ \vdots \\ x_n^m \end{bmatrix}}_{\mathbf{u}_m}$$

Portanto, $\mathbf{p} = \alpha_0^* \mathbf{u}_0 + \alpha_1^* \mathbf{u}_1 + \alpha_2^* \mathbf{u}_2 + \cdots + \alpha_m^* \mathbf{u}_m$.

Aproximação polinomial: caso discreto

Assim, podemos reescrever o vetor $\mathbf{p}$ da seguinte forma:

$$\mathbf{p} = \alpha_0^* \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{u}_0} + \alpha_1^* \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{u}_1} + \alpha_2^* \underbrace{\begin{bmatrix} x_0^2 \\ x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}}_{\mathbf{u}_2} + \cdots + \alpha_m^* \underbrace{\begin{bmatrix} x_0^m \\ x_1^m \\ \vdots \\ x_n^m \end{bmatrix}}_{\mathbf{u}_m}$$

Portanto, $\mathbf{p} = \alpha_0^* \mathbf{u}_0 + \alpha_1^* \mathbf{u}_1 + \alpha_2^* \mathbf{u}_2 + \cdots + \alpha_m^* \mathbf{u}_m$.

Afirmação: O conjunto $\mathcal{U} = \{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$ é L.I. (logo é uma base de $\mathbb{R}^{m+1}$, com $m < n$.)

Aproximação polinomial: caso discreto

Prova: Por hipótese temos que os os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ são distintos. Além disso, podemos escrever $\mathbf{p}$ na forma matricial:

Aproximação polinomial: caso discreto

Prova: Por hipótese temos que os os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ são distintos. Além disso, podemos escrever $\mathbf{p}$ na forma matricial:

$$\mathbf{p} = \underbrace{\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^m \\ 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}}_{\mathbf{X}} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix}$$

Aproximação polinomial: caso discreto

Prova: Por hipótese temos que os os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ são distintos. Além disso, podemos escrever $\mathbf{p}$ na forma matricial:

$$\mathbf{p} = \underbrace{\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^m \\ 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}}_{\mathbf{X}} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix}$$

Considere X' a submatriz quadrada de ordem $(m+1)$ de X:

$$\mathbf{X}' = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^m \\ 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & x_m & x_m^2 & \cdots & x_m^m \end{bmatrix}$$

Aproximação polinomial: caso discreto

Para mostrar que $\{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$ é L.I. Basta mostrar que $\mathbf{X}'$ é invertível ($\det(\mathbf{X}') \neq 0$).

Aproximação polinomial: caso discreto

Para mostrar que $\{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$ é L.I. Basta mostrar que $\mathbf{X}'$ é invertível ($\det(\mathbf{X}') \neq 0$).

Note que $\mathbf{X}'$ é uma *matriz de Vandermond*, logo:

$$\det(\mathbf{X}') = \prod_{i>j}^{m} (x_i - x_j)$$

Aproximação polinomial: caso discreto

Para mostrar que $\{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$ é L.I. Basta mostrar que $\mathbf{X}'$ é invertível ($\det(\mathbf{X}') \neq 0$).

Note que $\mathbf{X}'$ é uma *matriz de Vandermond*, logo:

$$\det(\mathbf{X}') = \prod_{i>j}^{m} (x_i - x_j)$$

Pelo fato de $x_i \neq x_j$ para $i \neq j$ implica que $\det(\mathbf{X}') \neq 0$. ■

Aproximação polinomial: caso discreto

Voltando ao problema de mínimos quadrados, queremos que a distância entre $\mathbf{y} \in \mathbb{R}^{n+1}$ e $\mathbf{p} \in \mathbb{R}^{n+1}$ seja mínima.

Aproximação polinomial: caso discreto

Voltando ao problema de mínimos quadrados, queremos que a distância entre $\mathbf{y} \in \mathbb{R}^{n+1}$ e $\mathbf{p} \in \mathbb{R}^{n+1}$ seja mínima.

Para isso, basta fazer a projeção ortogonal de $\mathbf{y}$ no subspaço gerado por $\mathcal{U} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$. Portanto, basta resolver o sistema de equações normais:

Aproximação polinomial: caso discreto

Voltando ao problema de mínimos quadrados, queremos que a distância entre $\mathbf{y} \in \mathbb{R}^{n+1}$ e $\mathbf{p} \in \mathbb{R}^{n+1}$ seja mínima.

Para isso, basta fazer a projeção ortogonal de $\mathbf{y}$ no subspaço gerado por $\mathcal{U} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^{n+1}$. Portanto, basta resolver o sistema de equações normais:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle \\ \langle \mathbf{u}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle \end{bmatrix}$$

onde $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{k=0}^{n} v_i w_i$.

Aproximação polinomial: caso discreto

Exemplo 19

Seja $f(x)$ uma função discreta dada pela tabela abaixo:

x	-1	0	1	2
$f(x)$	0	-1	0	7

Gostaríamos de aproximar no sentido de mínimos quadrados, a função $f(x)$ por uma parábola, ou seja:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$$

Aproximação polinomial: caso discreto

Exemplo 19

Seja $f(x)$ uma função discreta dada pela tabela abaixo:

x	-1	0	1	2
$f(x)$	0	-1	0	7

Gostaríamos de aproximar no sentido de mínimos quadrados, a função $f(x)$ por uma parábola, ou seja:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$$

Sabemos que o conjunto abaixo forma uma base de $\mathbb{R}^3$:

$$\mathcal{U} = \{\mathbf{u}_0 = (1,1,1,1)^\top, \mathbf{u}_1 = (-1,0,1,2)^\top, \mathbf{u}_2 = (1,0,1,4)^\top\},$$

Agora, vamos calcular a projeção ortogonal de $\mathbf{y} = (0,-1,0,7)^\top$ no subspaço gerado por $\mathcal{U}$, para isso devemos montar e resolver o sistema

Aproximação polinomial: caso discreto

Continuação

Mas antes, precisamos calcular os elementos da matriz **A** e do vetor **b**:

$$
\begin{array}{rclcrcl}
\langle u_0, u_0 \rangle &=& \sum_{k=0}^{3} 1 = 4 & & \langle u_0, u_1 \rangle &=& \sum_{k=0}^{3} x_k = 2 \\
\langle u_0, u_2 \rangle &=& \sum_{k=0}^{3} x_k^2 = 6 & & \langle u_1, u_0 \rangle &=& \langle u_0, u_1 \rangle = 2 \\
\langle u_1, u_1 \rangle &=& \langle u_0, u_2 \rangle = 6 & & \langle u_1, u_2 \rangle &=& \sum_{k=0}^{3} x_k^3 = 8 \\
\langle u_2, u_0 \rangle &=& \langle u_0, u_2 \rangle = 6 & & \langle u_2, u_1 \rangle &=& \langle u_1, u_2 \rangle = 8 \\
\langle u_2, u_2 \rangle &=& \sum_{k=0}^{3} x_k^4 = 18 & & \langle u_0, y \rangle &=& \sum_{k=0}^{3} y_k = 6 \\
\langle u_1, y \rangle &=& \sum_{k=0}^{3} x_k y_k = 14 & & \langle u_2, y \rangle &=& \sum_{k=0}^{3} x_k^2 y_k = 28
\end{array}
$$

Aproximação polinomial: caso discreto

Continuação

Finalmente,

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \alpha_2^* \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 28 \end{bmatrix}$$

cuja solução é $\alpha_0^* = -\frac{8}{5}$, $\alpha_1^* = \frac{1}{5}$ e $\alpha_2^* = 2$.

Portanto, a função procurada é

$$P_2(x) = -\frac{8}{5} + \frac{1}{5}x + 2x^2 \,.$$

Erro de Truncamento: caso discreto

Exemplo 20

Vamos calcular o erro de truncamento do exemplo anterior. Logo,

$$Q = \|f - P_2\|^2 \;=\; \langle f - P_2, f - P_2 \rangle = \sum_{k=0}^{3} (y_k - P_2(x_k))^2$$

Erro de Truncamento: caso discreto

Exemplo 20

Vamos calcular o erro de truncamento do exemplo anterior. Logo,

$$
\begin{aligned}
Q = \|f - P_2\|^2 &= \langle f - P_2, f - P_2 \rangle = \sum_{k=0}^{3} (y_k - P_2(x_k))^2 \\
&= \sum_{k=0}^{3} y_k^2 - 2\sum_{k=0}^{3} y_k P_2(x_k) + \sum_{k=0}^{3} P_2^2(x_k)
\end{aligned}
$$

Erro de Truncamento: caso discreto

Exemplo 20

Vamos calcular o erro de truncamento do exemplo anterior. Logo,

$$
\begin{aligned}
Q = \|f - P_2\|^2 &= \langle f - P_2, f - P_2 \rangle = \sum_{k=0}^{3} (y_k - P_2(x_k))^2 \\
&= \sum_{k=0}^{3} y_k^2 - 2\sum_{k=0}^{3} y_k P_2(x_k) + \sum_{k=0}^{3} P_2^2(x_k) \\
&= 50 - 98.4 + 49.2 = 0.8
\end{aligned}
$$

Aproximação polinomial: forma matricial

Dada uma função $f(x)$ amostrada , ou seja, é conhecida apenas nos $n+1$ pares de pontos $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, onde $y_i = f(x_i)$, $i = 0, \ldots, n$, com os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Vimos anteriormente que para calcular a seguinte aproximação no sentido de mínimos quadrados:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x), \text{ com } m < n$$

Aproximação polinomial: forma matricial

Dada uma função $f(x)$ amostrada , ou seja, é conhecida apenas nos $n+1$ pares de pontos $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, onde $y_i = f(x_i)$, $i = 0, \ldots, n$, com os $(n+1)$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Vimos anteriormente que para calcular a seguinte aproximação no sentido de mínimos quadrados:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x), \text{ com } m < n$$

Basta resolver o sistema linear:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle \\ \langle \mathbf{u}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle \end{bmatrix}$$

onde $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{k=0}^{n} v_i w_i$.

Aproximação polinomial: forma matricial

Entretanto, podemos reescrever o sistema linear:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle \\ \langle \mathbf{u}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle \end{bmatrix}$$

Da seguinte forma:

Aproximação polinomial: forma matricial

Entretanto, podemos reescrever o sistema linear:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle \\ \langle \mathbf{u}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle \end{bmatrix}$$

Da seguinte forma:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} 1 & x_0 & \cdots & x_0^m \\ 1 & x_1 & \cdots & x_1^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Aproximação polinomial: forma matricial

Entretanto, podemos reescrever o sistema linear:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle \\ \langle \mathbf{u}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle \end{bmatrix}$$

Da seguinte forma:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} 1 & x_0 & \cdots & x_0^m \\ 1 & x_1 & \cdots & x_1^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_0^m & x_1^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Na notação matricial, temos:

$$\mathbf{X}^\top \mathbf{X} \alpha^* = \mathbf{X}^\top \mathbf{y} \Rightarrow \alpha^* = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

MATLAB code

```
function a = mmq(x,y,k)

n = length(x);
X = vander(x);
X = X(:,n-k:n);
a = ((X'*X)\(X'*y'))';
```

Aproximação polinomial: forma matricial revisitada

Vimos anteriormente que o sistema de equações normais pode ser escrito na seguinte forma matricial:

$$\mathbf{X}^\top \mathbf{X} \alpha^* = \mathbf{X}^\top \mathbf{y}$$

Portanto, para determinar α^* basta resolver o sistema linear acima.

Aproximação polinomial: forma matricial revisitada

Vimos anteriormente que o sistema de equações normais pode ser escrito na seguinte forma matricial:

$$\mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y}$$

Portanto, para determinar α^* basta resolver o sistema linear acima. Partindo desse objetivo, vamos utilizar a *Decomposição QR* da matriz $\mathbf{X}$, isto é:

$$\mathbf{X} = \mathbf{QR}\,,$$

onde $\mathbf{Q}$ é uma matriz ortogonal de ordem $m \times m$ e $\mathbf{R}$ é uma matriz triangular superior de ordem $m \times n$. Logo,

$$\mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y} \quad \Leftrightarrow \quad (\mathbf{QR})^\top (\mathbf{QR})\alpha^* = (\mathbf{QR})^\top \mathbf{y} \tag{1}$$

Aproximação polinomial: forma matricial revisitada

Vimos anteriormente que o sistema de equações normais pode ser escrito na seguinte forma matricial:

$$\mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y}$$

Portanto, para determinar α^* basta resolver o sistema linear acima. Partindo desse objetivo, vamos utilizar a *Decomposição QR* da matriz $\mathbf{X}$, isto é:

$$\mathbf{X} = \mathbf{QR}\,,$$

onde $\mathbf{Q}$ é uma matriz ortogonal de ordem $m \times m$ e $\mathbf{R}$ é uma matriz triangular superior de ordem $m \times n$. Logo,

$$\begin{aligned} \mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y} \quad &\Leftrightarrow \quad (\mathbf{QR})^\top (\mathbf{QR})\alpha^* = (\mathbf{QR})^\top \mathbf{y} && (1) \\ &\Leftrightarrow \quad \mathbf{R}^\top (\mathbf{Q}^\top \mathbf{Q})\mathbf{R}\alpha^* = \mathbf{R}^\top \mathbf{Q}^\top \mathbf{y} && (2) \end{aligned}$$

Aproximação polinomial: forma matricial revisitada

Vimos anteriormente que o sistema de equações normais pode ser escrito na seguinte forma matricial:

$$\mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y}$$

Portanto, para determinar α^* basta resolver o sistema linear acima. Partindo desse objetivo, vamos utilizar a *Decomposição QR* da matriz $\mathbf{X}$, isto é:

$$\mathbf{X} = \mathbf{QR}\,,$$

onde $\mathbf{Q}$ é uma matriz ortogonal de ordem $m \times m$ e $\mathbf{R}$ é uma matriz triangular superior de ordem $m \times n$. Logo,

$$\begin{aligned}
\mathbf{X}^\top \mathbf{X}\alpha^* = \mathbf{X}^\top \mathbf{y} \quad &\Leftrightarrow \quad (\mathbf{QR})^\top (\mathbf{QR})\alpha^* = (\mathbf{QR})^\top \mathbf{y} && (1)\\
&\Leftrightarrow \quad \mathbf{R}^\top (\mathbf{Q}^\top \mathbf{Q})\mathbf{R}\alpha^* = \mathbf{R}^\top \mathbf{Q}^\top \mathbf{y} && (2)\\
&\Leftrightarrow \quad \mathbf{R}\alpha^* = \mathbf{Q}^\top \mathbf{y} && (3)
\end{aligned}$$

MATLAB code revisited

```
function a = mmq_qr(x,y,k)

n = length(x);
X = vander(x);
X = X(:,n-k:n);
[Q R] = qr(X);
a = (R\(Q'*y'))';
```

Impondo Restrições

No Exemplo 19, pelo método dos mínimos quadrados, obtemos a parábola

$$P_2(x) = -\frac{8}{5} + \frac{1}{5} + 2x^2 \, .$$

Impondo Restrições

No Exemplo 19, pelo método dos mínimos quadrados, obtemos a parábola

$$P_2(x) = -\frac{8}{5} + \frac{1}{5} + 2x^2 \,.$$

Agora queremos obter uma função $P_2(x)$ mais próxima de $f(x)$ no sentido dos mínimos quadrados dentre os polinômios de $\mathcal{P}_2$ que satisfazem uma restrição $P_2(2) = 7$. Para isso, precisamos do seguinte teorema.

Impondo Restrições

No Exemplo 19, pelo método dos mínimos quadrados, obtemos a parábola

$$P_2(x) = -\frac{8}{5} + \frac{1}{5} + 2x^2 .$$

Agora queremos obter uma função $P_2(x)$ mais próxima de $f(x)$ no sentido dos mínimos quadrados dentre os polinômios de $\mathcal{P}_2$ que satisfazem uma restrição $P_2(2) = 7$. Para isso, precisamos do seguinte teorema.

Teorema

F^* é a melhor aproximação de f dentre as combinações lineares $\alpha_0\phi_0 + \alpha_1\phi_1 + \cdots + \alpha_n\phi_n$ se somente se G^* é a melhor aproximação de $f + \psi$ dentre as combinações lineares $\alpha_0\phi_0 + \alpha_1\phi_1 + \cdots + \alpha_n\phi_n + \psi$.

Impondo Restrições

Logo, queremos $f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$ sob restrição

$$\alpha_0^* + 2\alpha_1^* + 4\alpha_2^* = 7\,.$$

Impondo Restrições

Logo, queremos $f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$ sob restrição

$$\alpha_0^* + 2\alpha_1^* + 4\alpha_2^* = 7\,.$$

1. Eliminar coeficiente

$$\alpha_2^* = \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} \implies P_2(x) = \alpha_0^* + \alpha_1^* x + \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} x^2$$

Impondo Restrições

Logo, queremos $f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$ sob restrição

$$\alpha_0^* + 2\alpha_1^* + 4\alpha_2^* = 7\,.$$

1. Eliminar coeficiente

$$\alpha_2^* = \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} \implies P_2(x) = \alpha_0^* + \alpha_1^* x + \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} x^2$$

2. Isolar coeficientes

$$f(x) \approx \underbrace{\frac{7}{4}x^2}_{-\psi} + \alpha_0^* \underbrace{\left(1 - \frac{1}{4}x^2\right)}_{\phi_0} + \alpha_1^* \underbrace{\left(x - \frac{1}{2}x^2\right)}_{\phi_1}$$

Impondo Restrições

Logo, queremos $f(x) \approx \alpha_0^* + \alpha_1^* x + \alpha_2^* x^2 = P_2(x)$ sob restrição

$$\alpha_0^* + 2\alpha_1^* + 4\alpha_2^* = 7\,.$$

1. Eliminar coeficiente

$$\alpha_2^* = \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} \implies P_2(x) = \alpha_0^* + \alpha_1^* x + \frac{7 - \alpha_0^* - 2\alpha_1^*}{4} x^2$$

2. Isolar coeficientes

$$f(x) \approx \underbrace{\frac{7}{4}x^2}_{-\psi} + \alpha_0^* \underbrace{\left(1 - \frac{1}{4}x^2\right)}_{\phi_0} + \alpha_1^* \underbrace{\left(x - \frac{1}{2}x^2\right)}_{\phi_1}$$

3. Transformar

$$\underbrace{f(x) - \frac{7}{4}x^2}_{G=f+\psi} \approx \alpha_0^* \underbrace{\left(1 - \frac{1}{4}x^2\right)}_{\phi_0} + \alpha_1^* \underbrace{\left(x - \frac{1}{2}x^2\right)}_{\phi_1}$$

Impondo Restrições

4. Identificar a base

$$\phi_0(x) = 1 - \frac{1}{4}x^2 \qquad \phi_1(x) = x - \frac{1}{2}x^2 \qquad dim(V) = 2\,.$$

Impondo Restrições

4. Identificar a base

$$\phi_0(x) = 1 - \frac{1}{4}x^2 \qquad \phi_1(x) = x - \frac{1}{2}x^2 \qquad dim(V) = 2\,.$$

5. Resolver o sistema linear

$$\begin{bmatrix} \langle \phi_0, \phi_0 \rangle & \langle \phi_0, \phi_1 \rangle \\ \langle \phi_1, \phi_0 \rangle & \langle \phi_1, \phi_1 \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} \langle G, \phi_0 \rangle \\ \langle G, \phi_1 \rangle \end{bmatrix} \Longrightarrow \begin{array}{rl} \alpha_0^* = & -1.6316 \\ \alpha_1^* = & 0.2105 \end{array}$$

Impondo Restrições

4. Identificar a base

$$\phi_0(x) = 1 - \frac{1}{4}x^2 \qquad \phi_1(x) = x - \frac{1}{2}x^2 \qquad dim(V) = 2\,.$$

5. Resolver o sistema linear

$$\begin{bmatrix} \langle \phi_0, \phi_0 \rangle & \langle \phi_0, \phi_1 \rangle \\ \langle \phi_1, \phi_0 \rangle & \langle \phi_1, \phi_1 \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} \langle G, \phi_0 \rangle \\ \langle G, \phi_1 \rangle \end{bmatrix} \Longrightarrow \begin{array}{rr} \alpha_0^* = & -1.6316 \\ \alpha_1^* = & 0.2105 \end{array}$$

6. Solução

$$P_2(x) = \frac{7}{4}x^2 - 1.6316\left(1 - \frac{1}{4}x^2\right) + 0.2105\left(x - \frac{1}{2}x^2\right)$$

Outras Aproximações

Vimos que o objetivo do método dos mínimos quadrados é aproximar uma função dada $f(x)$ por uma combinação linear:

$$F^*(x) = \alpha_0^*\phi_0 + \alpha_1^*\phi_1 + \cdots + \alpha_n^*\phi_n$$

Outras Aproximações

Vimos que o objetivo do método dos mínimos quadrados é aproximar uma função dada $f(x)$ por uma combinação linear:

$$F^*(x) = \alpha_0^*\phi_0 + \alpha_1^*\phi_1 + \cdots + \alpha_n^*\phi_n$$

Mas como aproximar $f(x)$ por uma função não-linear?

Outras Aproximações

Vimos que o objetivo do método dos mínimos quadrados é aproximar uma função dada $f(x)$ por uma combinação linear:

$$F^*(x) = \alpha_0^*\phi_0 + \alpha_1^*\phi_1 + \cdots + \alpha_n^*\phi_n$$

Mas como aproximar $f(x)$ por uma função não-linear? Simples, basta linearizar o problema!

Outras Aproximações

Vimos que o objetivo do método dos mínimos quadrados é aproximar uma função dada $f(x)$ por uma combinação linear:

$$F^*(x) = \alpha_0^*\phi_0 + \alpha_1^*\phi_1 + \cdots + \alpha_n^*\phi_n$$

Mas como aproximar $f(x)$ por uma função não-linear? Simples, basta linearizar o problema!

1. Caso $f(x) \approx a\, b^x$:

$$\underbrace{\ln\,(f(x))}_{F(x)} \approx \underbrace{\ln\,(a)}_{\alpha_0^*}\,\underbrace{1}_{\phi_0} + \underbrace{\ln\,(b)}_{\alpha_1^*}\,\underbrace{x}_{\phi_1}\,.$$

Após resolver o sistema linear, temos $a = e^{\alpha_0^*}$ e $b = e^{\alpha_1^*}$.

Outras Aproximações

Vimos que o objetivo do método dos mínimos quadrados é aproximar uma função dada $f(x)$ por uma combinação linear:

$$F^*(x) = \alpha_0^*\phi_0 + \alpha_1^*\phi_1 + \cdots + \alpha_n^*\phi_n$$

Mas como aproximar $f(x)$ por uma função não-linear? Simples, basta linearizar o problema!

1. Caso $f(x) \approx a\,b^x$:

$$\underbrace{\ln\,(f(x))}_{F(x)} \approx \underbrace{\ln\,(a)}_{\alpha_0^*}\,\underbrace{1}_{\phi_0} + \underbrace{\ln\,(b)}_{\alpha_1^*}\,\underbrace{x}_{\phi_1}.$$

   Após resolver o sistema linear, temos $a = e^{\alpha_0^*}$ e $b = e^{\alpha_1^*}$.

2. Caso $f(x) \approx a\,x^b$:

$$\underbrace{\ln\,(f(x))}_{F(x)} \approx \underbrace{\ln\,(a)}_{\alpha_0^*}\,\underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*}\,\underbrace{\ln\,(x)}_{\phi_1}.$$

Outras Aproximações

Outras Aproximações

3 Caso $f(x) \approx \frac{1}{a+bx}$:

$$\underbrace{\frac{1}{f(x)}}_{F(x)} \approx \underbrace{a}_{\alpha_0^*} \underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*} \underbrace{x}_{\phi_1}$$

Outras Aproximações

3 Caso $f(x) \approx \frac{1}{a+bx}$:

$$\underbrace{\frac{1}{f(x)}}_{F(x)} \approx \underbrace{a}_{\alpha_0^*}\underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*}\underbrace{x}_{\phi_1}$$

4 Caso $f(x) \approx \sqrt{a+bx}$:

$$\underbrace{[f(x)]^2}_{F(x)} \approx \underbrace{a}_{\alpha_0^*}\underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*}\underbrace{x}_{\phi_1}$$

Outras Aproximações

3. Caso $f(x) \approx \frac{1}{a+bx}$:

$$\underbrace{\frac{1}{f(x)}}_{F(x)} \approx \underbrace{a}_{\alpha_0^*} \underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*} \underbrace{x}_{\phi_1}$$

4. Caso $f(x) \approx \sqrt{a+bx}$:

$$\underbrace{[f(x)]^2}_{F(x)} \approx \underbrace{a}_{\alpha_0^*} \underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*} \underbrace{x}_{\phi_1}$$

5. Caso $f(x) \approx x \ln(a+bx)$:

$$\underbrace{e^{\frac{f(x)}{x}}}_{F(x)} \approx \underbrace{a}_{\alpha_0^*} \underbrace{1}_{\phi_0} + \underbrace{b}_{\alpha_1^*} \underbrace{x}_{\phi_1}$$

Outras Aproximações

Exemplo 21

Seja $y = f(x)$ uma função dada pela tabela abaixo:

x	0	1	2	3
y	1	1	1.7	2.5

Aproxime $f(x)$, no sentido de mínimos quadrados, por uma função racional do tipo:

$$f(x) \approx \frac{a + x^2}{b + x}.$$

Outras Aproximações

Exemplo 21

Seja $y = f(x)$ uma função dada pela tabela abaixo:

x	0	1	2	3
y	1	1	1.7	2.5

Aproxime $f(x)$, no sentido de mínimos quadrados, por uma função racional do tipo:

$$f(x) \approx \frac{a + x^2}{b + x}.$$

Solução:

$$\underbrace{f(x)}_{F(x)} \approx \underbrace{\frac{a}{b}}_{\alpha_0^*} \underbrace{1}_{\phi_0} + \underbrace{\frac{1}{b}}_{\alpha_1^*} \underbrace{x\left[x - f(x)\right]}_{\phi_1}$$

Outras Aproximações

Solução (continuação): Assim, $\phi_0 = (1,1,1,1)^\top$, $\phi_1 = (0,0,0.6,1.5)^\top$ e $y = (1,1,1.7,2.5)^\top$. Logo,

$$\begin{bmatrix} \langle \phi_0, \phi_0 \rangle & \langle \phi_0, \phi_1 \rangle \\ \langle \phi_0, \phi_1 \rangle & \langle \phi_1, \phi_1 \rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} \langle \phi_0, y \rangle \\ \langle \phi_1, y \rangle \end{bmatrix} \Rightarrow \begin{bmatrix} 4.00 & 2.10 \\ 2.10 & 2.61 \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} 6.20 \\ 4.77 \end{bmatrix}$$

Outras Aproximações

Solução (continuação): Assim, $\phi_0 = (1,1,1,1)^\top$, $\phi_1 = (0,0,0.6,1.5)^\top$ e $y = (1,1,1.7,2.5)^\top$. Logo,

$$\begin{bmatrix} \langle\phi_0,\phi_0\rangle & \langle\phi_0,\phi_1\rangle \\ \langle\phi_0,\phi_1\rangle & \langle\phi_1,\phi_1\rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} \langle\phi_0,y\rangle \\ \langle\phi_1,y\rangle \end{bmatrix} \Rightarrow \begin{bmatrix} 4.00 & 2.10 \\ 2.10 & 2.61 \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} 6.20 \\ 4.77 \end{bmatrix}$$

Cuja solução é $\alpha_0^* = 1.0224$ e $\alpha_1^* = 1.0050$.

Outras Aproximações

Solução (continuação): Assim, $\phi_0 = (1,1,1,1)^\top$, $\phi_1 = (0,0,0.6,1.5)^\top$ e $y = (1,1,1.7,2.5)^\top$. Logo,

$$\begin{bmatrix} \langle\phi_0,\phi_0\rangle & \langle\phi_0,\phi_1\rangle \\ \langle\phi_0,\phi_1\rangle & \langle\phi_1,\phi_1\rangle \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} \langle\phi_0,y\rangle \\ \langle\phi_1,y\rangle \end{bmatrix} \Rightarrow \begin{bmatrix} 4.00 & 2.10 \\ 2.10 & 2.61 \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \end{bmatrix} = \begin{bmatrix} 6.20 \\ 4.77 \end{bmatrix}$$

Cuja solução é $\alpha_0^* = 1.0224$ e $\alpha_1^* = 1.0050$.

Portanto,

$$b = \frac{1}{\alpha_1^*} = 0.9950 \qquad a = \alpha_0^*\, b = 0.0173\,.$$

$$f(x) \approx \frac{0.0173 + x^2}{0.9950 + x}\,.$$

Produto escalar com pesos

Dados dois vetores $\mathbf{u} = (u_0, u_1, \ldots, u_n)^\top$ e $\mathbf{v} = (v_0, v_1, \ldots, v_n)^\top$ de $V = \mathbb{R}^{n+1}$. O produto interno usual para V é dado por:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=0}^{n} u_i v_i, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n+1}$$

Seja $\mathbf{w} = (w_0, w_1, \ldots, w_n)^\top$, com $w_i > 0, \ \forall i = 0, \ldots, n$.

Será que a aplicação

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum_{i=0}^{n} w_i u_i v_i$$

define um produto interno em $\mathbb{R}^{n+1}$?

Produto escalar com pesos

Vejamos:

- bilinearidade?

$$
\begin{aligned}
\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{z} \rangle_{\mathbf{w}} &= \sum w_i u_i(\alpha v_i + \beta z_i) = \sum \alpha w_i u_i v_i + \beta w_i u_i z_i \\
&= \alpha \sum w_i u_i v_i + \beta \sum w_i u_i z_i = \alpha \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} + \beta \langle \mathbf{u}, \mathbf{z} \rangle_{\mathbf{w}}
\end{aligned}
$$

Produto escalar com pesos

Vejamos:

- bilinearidade?

$$
\begin{aligned}
\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{z} \rangle_{\mathbf{w}} &= \sum w_i u_i(\alpha v_i + \beta z_i) = \sum \alpha w_i u_i v_i + \beta w_i u_i z_i \\
&= \alpha \sum w_i u_i v_i + \beta \sum w_i u_i z_i = \alpha \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} + \beta \langle \mathbf{u}, \mathbf{z} \rangle_{\mathbf{w}}
\end{aligned}
$$

- simetria?

$$
\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum w_i u_i v_i = \sum w_i v_i u_i = \langle \mathbf{v}, \mathbf{u} \rangle_{\mathbf{w}}
$$

Produto escalar com pesos

Vejamos:

- bilinearidade?

$$\begin{aligned} \langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{z} \rangle_{\mathbf{w}} &= \sum w_i u_i(\alpha v_i + \beta z_i) = \sum \alpha w_i u_i v_i + \beta w_i u_i z_i \\ &= \alpha \sum w_i u_i v_i + \beta \sum w_i u_i z_i = \alpha \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} + \beta \langle \mathbf{u}, \mathbf{z} \rangle_{\mathbf{w}} \end{aligned}$$

- simetria?

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum w_i u_i v_i = \sum w_i v_i u_i = \langle \mathbf{v}, \mathbf{u} \rangle_{\mathbf{w}}$$

- positividade?

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbf{w}} = \sum w_i u_i^2 \geq 0, \quad \text{já que } w_i > 0 \text{ e } u_i^2 \geq 0$$

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbf{w}} = 0 \quad \Leftrightarrow \quad \sum w_i u_i^2 = 0 \quad \Leftrightarrow \quad u_i = 0, \ \forall\, i \quad \Leftrightarrow \quad \mathbf{u} = \overline{0}$$

Mínimos quadrados ponderados

Desta forma, podemos adotar o produto

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum_{i=0}^{n} w_i u_i v_i$$

como produto escalar em $\mathbb{R}^{n+1}$.

Mínimos quadrados ponderados

Desta forma, podemos adotar o produto

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum_{i=0}^{n} w_i u_i v_i$$

como produto escalar em $\mathbb{R}^{n+1}$.

Utilizando o produto interno acima, podemos calcular a aproximação por mínimos quadrados de uma função discreta:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n), \quad \text{com} \quad y_i = f(x_i), \forall i$$

por uma função em um espaço de dimensão $m + 1 < n + 1$.

Mínimos quadrados ponderados

Desta forma, podemos adotar o produto

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \sum_{i=0}^{n} w_i u_i v_i$$

como produto escalar em $\mathbb{R}^{n+1}$.

Utilizando o produto interno acima, podemos calcular a aproximação por mínimos quadrados de uma função discreta:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n), \quad \text{com} \quad y_i = f(x_i), \, \forall i$$

por uma função em um espaço de dimensão $m + 1 < n + 1$.

O procedimento é o mesmo que o anterior, trocando-se $\langle \cdot, \cdot \rangle$ por $\langle \cdot, \cdot \rangle_{\mathbf{w}}$. Assim, desejamos obter a função $F^* = \alpha_0^* \varphi_0 + \alpha_1^* \varphi_1 + \cdots + \alpha_m^* x^m \varphi_m$ que melhor aproxima f sob a distância:

$$dist(f, F^*)_{\mathbf{w}} = \|f - F^*\|_{\mathbf{w}} = \sqrt{\langle f - F^*, f - F^* \rangle_{\mathbf{w}}}$$

Mínimos quadrados ponderados

De maneira análoga, seja V um subespaço vetorial de $\mathcal{C}[a,b]$, de dimensão finita.

Seja ainda $w(x)$ uma função contínua, com $w(x) > 0, \forall x \in [a,b]$.

Desta forma

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

define um produto escalar em V.

Aproximação polinomial: caso discreto

Dados os pesos $\mathbf{w} = (w_0, w_1, \ldots, w_n)^\top$ e uma função $f(x)$ amostrada:

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n), \quad \text{com} \quad y_i = f(x_i), \forall i$$

Desejamos calcular a seguinte aproximação no sentido de mínimos quadrados ponderados:

$$f(x) \approx \alpha_0^* + \alpha_1^* x + \cdots + \alpha_m^* x^m = P_m(x), \text{ com } m < n$$

Aproximação polinomial: forma matricial

Para isso precisamos resolver o sistema linear abaixo:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \vdots & \vdots & \ldots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle_{\mathbf{w}} \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{y} \rangle_{\mathbf{w}} \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle_{\mathbf{w}} \end{bmatrix}$$

Aproximação polinomial: forma matricial

Para isso precisamos resolver o sistema linear abaixo:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \vdots & \vdots & \ldots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle_{\mathbf{w}} \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{y} \rangle_{\mathbf{w}} \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle_{\mathbf{w}} \end{bmatrix}$$

Note que o sistema acima pode ser escrito da seguinte maneira:

$$\mathbf{M}\alpha^* = \mathbf{X}^\top \mathbf{W} \mathbf{y}$$

onde

$$\mathbf{M} = \mathbf{X}^\top \mathbf{W} \mathbf{X} \quad \text{e} \quad \mathbf{W} = \operatorname{diag}(w_0, w_1, \ldots, w_n)$$

Aproximação polinomial: forma matricial

Para isso precisamos resolver o sistema linear abaixo:

$$\begin{bmatrix} \langle \mathbf{u}_0, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_0, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_0, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_1, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_1, \mathbf{u}_m \rangle_{\mathbf{w}} \\ \vdots & \vdots & \ldots & \vdots \\ \langle \mathbf{u}_m, \mathbf{u}_0 \rangle_{\mathbf{w}} & \langle \mathbf{u}_m, \mathbf{u}_1 \rangle_{\mathbf{w}} & \cdots & \langle \mathbf{u}_m, \mathbf{u}_m \rangle_{\mathbf{w}} \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \vdots \\ \alpha_m^* \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_0, \mathbf{y} \rangle_{\mathbf{w}} \\ \langle \mathbf{u}_1, \mathbf{y} \rangle_{\mathbf{w}} \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{y} \rangle_{\mathbf{w}} \end{bmatrix}$$

Note que o sistema acima pode ser escrito da seguinte maneira:

$$\mathbf{M}\alpha^* = \mathbf{X}^\top \mathbf{W} \mathbf{y}$$

onde

$$\mathbf{M} = \mathbf{X}^\top \mathbf{W} \mathbf{X} \quad \text{e} \quad \mathbf{W} = \operatorname{diag}(w_0, w_1, \ldots, w_n)$$

Portanto, a solução do sistema é dada por:

$$\alpha^* = \mathbf{M}^{-1} \mathbf{X}^\top \mathbf{W} \mathbf{y}$$

Aproximação polinomial

Exemplo

Considere a função discreta do exemplo anterior. Vamos definir os seguintes pesos $\mathbf{w} = (w_0, w_1, w_2, w_3) = (0.2, 1, 1, 0.2)$. Desta forma podemos calcular os produtos escalares

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = 0.2\, u_1 v_1 + u_2 v_2 + u_3 v_3 + 0.2\, u_4 v_4$$

resultando no sistema

$$\begin{bmatrix} 2.4 & 1.2 & 2 \\ 1.2 & 2 & 2.4 \\ 2 & 2.4 & 4.4 \end{bmatrix} \begin{bmatrix} \alpha_0^* \\ \alpha_1^* \\ \alpha_2^* \end{bmatrix} = \begin{bmatrix} 0.4 \\ 2.8 \\ 5.6 \end{bmatrix}$$

cuja solução é $\alpha_0^* = -1.4286$, $\alpha_1^* = -0.1429$ e $\alpha_2^* = 2$. Portanto,

$$P_2(x) = -1.4286 - 0.1429x + 2x^2 \,.$$

MATLAB code

```
function a = mmqp(x,y,w,k)
%Weighted-least squares

n = length(x);
W = diag(w);
X = vander(x);
X = X(:,n-k:n);
a = ((X'*W*X)\(X'*W*y'))';
```

Aproximação trigonométrica: caso contínuo

Dada uma função $f \in \mathcal{C}([-\pi, \pi])$ desejamos aproximar $f(x)$ no sentido de mínimos quadrados por um polinômio trigonométrico $S_n(x) \in \mathcal{T}_n$ na forma:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \operatorname{sen}(kx)) = S_n(x)$$

Aproximação trigonométrica: caso contínuo

Dada uma função $f \in \mathcal{C}([-\pi,\pi])$ desejamos aproximar $f(x)$ no sentido de mínimos quadrados por um polinômio trigonométrico $S_n(x) \in \mathcal{T}_n$ na forma:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \operatorname{sen}(kx)) = S_n(x)$$

Para isso, vamos calcular a projeção ortogonal de $f(x)$ no subspaço das funções trigonométricas $\mathcal{T}_n$ cuja uma base é dada por:

$$\mathcal{B} = \left\{ \frac{1}{2}, \cos(x), \operatorname{sen}(x), \cos(2x), \operatorname{sen}(2x), \ldots, \cos(nx), \operatorname{sen}(nx) \right\}$$

Aproximação trigonométrica: caso contínuo

Dada uma função $f \in \mathcal{C}([-\pi, \pi])$ desejamos aproximar $f(x)$ no sentido de mínimos quadrados por um polinômio trigonométrico $S_n(x) \in \mathcal{T}_n$ na forma:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \operatorname{sen}(kx)) = S_n(x)$$

Para isso, vamos calcular a projeção ortogonal de $f(x)$ no subspaço das funções trigonométricas $\mathcal{T}_n$ cuja uma base é dada por:

$$\mathcal{B} = \left\{ \frac{1}{2}, \cos(x), \operatorname{sen}(x), \cos(2x), \operatorname{sen}(2x), \ldots, \cos(nx), \operatorname{sen}(nx) \right\}$$

O limite de $S_n(x)$, quando $n \to \infty$, é chamado de *série de Fourier* de $f(x)$.

Aproximação trigonométrica: caso contínuo

Dessa forma para obter $S_n(x)$ que minimiza $Q = \|f(x) - S_n(x)\|^2$, basta determinar os coeficientes $a_0, a_1, b_1, \ldots, a_n, b_n$, resolvendo o sistema linear:

Aproximação trigonométrica: caso contínuo

Dessa forma para obter $S_n(x)$ que minimiza $Q = \|f(x) - S_n(x)\|^2$, basta determinar os coeficientes $a_0, a_1, b_1, \ldots, a_n, b_n$, resolvendo o sistema linear:

$$\begin{bmatrix} \left\langle \frac{1}{2}, \frac{1}{2} \right\rangle & \left\langle \frac{1}{2}, \cos(x) \right\rangle & \cdots & \left\langle \frac{1}{2}, \operatorname{sen}(nx) \right\rangle \\ \left\langle \cos(x), \frac{1}{2} \right\rangle & \langle \cos(x), \cos(x) \rangle & \cdots & \langle \cos(x), \operatorname{sen}(nx) \rangle \\ \left\langle \operatorname{sen}(x), \frac{1}{2} \right\rangle & \langle \operatorname{sen}(x), \cos(x) \rangle & \cdots & \langle \operatorname{sen}(x), \operatorname{sen}(nx) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle \operatorname{sen}(nx), \frac{1}{2} \right\rangle & \langle \operatorname{sen}(nx), \cos(x) \rangle & \cdots & \langle \operatorname{sen}(x), \operatorname{sen}(nx) \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \left\langle \frac{1}{2}, f \right\rangle \\ \langle \cos(x), f \rangle \\ \langle \operatorname{sen}(x), f \rangle \\ \vdots \\ \langle \operatorname{sen}(nx), f \rangle \end{bmatrix}$$

onde

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x) g(x) dx.$$

Aproximação trigonométrica: caso contínuo

Por outro lado, usando o fato que:

$$\cos(a \pm b) = \cos(a)\cos(b) \mp \operatorname{sen}(a)\operatorname{sen}(b)$$

$$\operatorname{sen}(a \pm b) = \operatorname{sen}(a)\cos(b) \pm \operatorname{sen}(b)\cos(a)$$

obtemos as seguintes identidades trigonométricas:

$$\operatorname{sen}(a)\operatorname{sen}(b) = \frac{1}{2}\left(\cos(a-b) - \cos(a+b)\right)$$

$$\cos(a)\cos(b) = \frac{1}{2}\left(\cos(a-b) + \cos(a+b)\right)$$

$$\operatorname{sen}(a)\cos(b) = \frac{1}{2}\left(\operatorname{sen}(a-b) + \operatorname{sen}(a+b)\right)$$

Aproximação trigonométrica: caso contínuo

Utilizando essas identidades segue que:

$$\langle \operatorname{sen}(px), \cos(qx) \rangle = \int_{-\pi}^{\pi} \operatorname{sen}(px) \cos(qx) dx = 0, \ \forall p, q \in \mathbb{N}$$

$$\langle \operatorname{sen}(px), \operatorname{sen}(qx) \rangle = \int_{-\pi}^{\pi} \operatorname{sen}(px) \operatorname{sen}(qx) dx = \begin{cases} 0 & p \neq q \\ \pi & p = q \neq 0 \end{cases}$$

$$\langle \cos(px), \cos(qx) \rangle = \int_{-\pi}^{\pi} \cos(px) \cos(qx) dx = \begin{cases} 0 & p \neq q \\ \pi & p = q \neq 0 \end{cases}$$

Aproximação trigonométrica: caso contínuo

Utilizando essas identidades segue que:

$$\langle \operatorname{sen}(px), \cos(qx)\rangle = \int_{-\pi}^{\pi} \operatorname{sen}(px)\cos(qx)dx = 0,\ \forall p, q \in \mathbb{N}$$

$$\langle \operatorname{sen}(px), \operatorname{sen}(qx)\rangle = \int_{-\pi}^{\pi} \operatorname{sen}(px)\operatorname{sen}(qx)dx = \begin{cases} 0 & p \neq q \\ \pi & p = q \neq 0 \end{cases}$$

$$\langle \cos(px), \cos(qx)\rangle = \int_{-\pi}^{\pi} \cos(px)\cos(qx)dx = \begin{cases} 0 & p \neq q \\ \pi & p = q \neq 0 \end{cases}$$

Logo, $\mathcal{B} = \left\{\frac{1}{2}, \cos(x), \operatorname{sen}(x), \cos(2x), \operatorname{sen}(2x), \ldots, \cos(nx), \operatorname{sen}(nx)\right\}$ é ortogonal em $[-\pi, \pi]$.

Aproximação trigonométrica: caso contínuo

Dessa maneira podemos reescrever o sistema de equações normais da seguinte forma:

Aproximação trigonométrica: caso contínuo

Dessa maneira podemos reescrever o sistema de equações normais da seguinte forma:

$$\begin{bmatrix} \frac{\pi}{2} & 0 & \cdots & 0 \\ 0 & \pi & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \pi \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \left\langle \frac{1}{2}, f \right\rangle \\ \langle \cos(x), f \rangle \\ \vdots \\ \langle \operatorname{sen}(nx), f \rangle \end{bmatrix}$$

Aproximação trigonométrica: caso contínuo

Dessa maneira podemos reescrever o sistema de equações normais da seguinte forma:

$$\begin{bmatrix} \frac{\pi}{2} & 0 & \cdots & 0 \\ 0 & \pi & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \pi \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \langle \frac{1}{2}, f \rangle \\ \langle \cos(x), f \rangle \\ \vdots \\ \langle \operatorname{sen}(nx), f \rangle \end{bmatrix}$$

Portanto,

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx, \quad k = 1, \ldots, n \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen}(kx) dx, \quad k = 1, \ldots, n \end{aligned}$$

Aproximação trigonométrica: caso contínuo

Considerações

1. Se a função dada $f(x)$ é par, como $\operatorname{sen}(kx)$ é uma função ímpar, então:

Aproximação trigonométrica: caso contínuo

Considerações

1. Se a função dada $f(x)$ é par, como $\operatorname{sen}(kx)$ é uma função ímpar, então:

$$\int_{-\pi}^{\pi} f(x)\operatorname{sen}(kx)dx = 0 \Rightarrow f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

Aproximação trigonométrica: caso contínuo

Considerações

1. Se a função dada $f(x)$ é par, como $\operatorname{sen}(kx)$ é uma função ímpar, então:

$$\int_{-\pi}^{\pi} f(x)\operatorname{sen}(kx)dx = 0 \;\Rightarrow\; f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

2. Se a função dada $f(x)$ é ímpar, como $\cos(kx)$ é uma função par, então:

Aproximação trigonométrica: caso contínuo

Considerações

1. Se a função dada $f(x)$ é par, como $\operatorname{sen}(kx)$ é uma função ímpar, então:

$$\int_{-\pi}^{\pi} f(x)\operatorname{sen}(kx)dx = 0 \Rightarrow f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

2. Se a função dada $f(x)$ é ímpar, como $\cos(kx)$ é uma função par, então:

$$\int_{-\pi}^{\pi} f(x)\cos(kx)dx = 0 \Rightarrow f(x) \approx \sum_{k=1}^{n} b_k \operatorname{sen}(kx) = S_n(x)$$

Aproximação trigonométrica: caso contínuo

Exemplo

Vamos calcular uma aproximação trigonométrica de ordem n para $f(x) = |x|$, para $x \in [-\pi, \pi]$.

Aproximação trigonométrica: caso contínuo

Exemplo

Vamos calcular uma aproximação trigonométrica de ordem n para $f(x) = |x|$, para $x \in [-\pi, \pi]$.

Note que $f(x)$ é uma função par, logo:

Aproximação trigonométrica: caso contínuo

Exemplo

Vamos calcular uma aproximação trigonométrica de ordem n para $f(x) = |x|$, para $x \in [-\pi, \pi]$.

Note que $f(x)$ é uma função par, logo:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

Aproximação trigonométrica: caso contínuo

Exemplo

Vamos calcular uma aproximação trigonométrica de ordem n para $f(x) = |x|$, para $x \in [-\pi, \pi]$.

Note que $f(x)$ é uma função par, logo:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

Assim,

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} |x| dx = -\frac{1}{\pi}\int_{-\pi}^{0} x dx + \frac{1}{\pi}\int_{0}^{\pi} x dx = \frac{2}{\pi}\int_{0}^{\pi} x dx = \pi$$

Aproximação trigonométrica: caso contínuo

Exemplo

Vamos calcular uma aproximação trigonométrica de ordem n para $f(x) = |x|$, para $x \in [-\pi, \pi]$.

Note que $f(x)$ é uma função par, logo:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) = S_n(x)$$

Assim,

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} |x| dx = -\frac{1}{\pi}\int_{-\pi}^{0} x dx + \frac{1}{\pi}\int_{0}^{\pi} x dx = \frac{2}{\pi}\int_{0}^{\pi} x dx = \pi$$

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} |x| \cos(kx) dx = \frac{2}{\pi}\int_{0}^{\pi} x \cos(kx) dx = \frac{2}{k^2\pi}\left((-1)^k - 1\right)$$

Aproximação trigonométrica: caso contínuo

Exemplo (continuação)

Portanto,

$$S_n(x) = \frac{\pi}{2} + \frac{2}{\pi} \sum_{k=1}^{n} \frac{(-1)^k - 1}{k^2} \cos(kx)$$

Aproximação trigonométrica: caso contínuo

Exemplo (continuação)

Portanto,

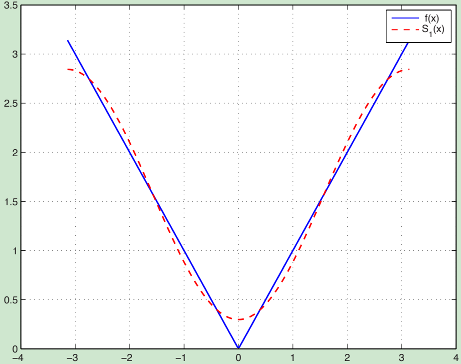
$$S_n(x) = \frac{\pi}{2} + \frac{2}{\pi} \sum_{k=1}^{n} \frac{(-1)^k - 1}{k^2} \cos(kx)$$

Além disso, a série de Fourier de $f(x)$ é:

$$S(x) = \lim_{n \to \infty} S_n(x) = \frac{\pi}{2} + \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^k - 1}{k^2} \cos(kx)$$

Aproximação trigonométrica: caso contínuo

Exemplo (continuação)



Aproximação trigonométrica de $f(x) = |x|$ por $S_1(x) = \frac{\pi}{2} - \frac{4}{\pi}\cos(x)$.

Aproximação trigonométrica: caso discreto

Dada uma função $f(x)$ conhecida apenas nos $2N$ pontos distintos $x_k = -\pi + \frac{k}{N}\pi$, com $k = 0, \ldots, 2N-1$. Desejamos aproximar $f(x)$ por $S_n(x) \in \mathcal{T}_n$, com $n \leq N$, no sentido de mínimos quadrados, isto é:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \operatorname{sen}(kx)\right) = S_n(x)\,,$$

tal que $Q = \min \|f(x) - S_n(x)\|^2$.

Aproximação trigonométrica: caso discreto

Dada uma função $f(x)$ conhecida apenas nos $2N$ pontos distintos $x_k = -\pi + \frac{k}{N}\pi$, com $k = 0, \ldots, 2N-1$. Desejamos aproximar $f(x)$ por $S_n(x) \in \mathcal{T}_n$, com $n \leq N$, no sentido de mínimos quadrados, isto é:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \operatorname{sen}(kx)\right) = S_n(x)\,,$$

tal que $Q = \min \|f(x) - S_n(x)\|^2$.

Para isso, vamos calcular a projeção ortogonal de $f(x)$ no subspaço gerado por:

$$\mathcal{B} = \left\{\frac{1}{2}, \cos(x), \operatorname{sen}(x), \cos(2x), \operatorname{sen}(2x), \ldots, \cos(nx), \operatorname{sen}(nx)\right\}.$$

Aproximação trigonométrica: caso discreto

Dada uma função $f(x)$ conhecida apenas nos $2N$ pontos distintos $x_k = -\pi + \frac{k}{N}\pi$, com $k = 0, \ldots, 2N-1$. Desejamos aproximar $f(x)$ por $S_n(x) \in \mathcal{T}_n$, com $n \leq N$, no sentido de mínimos quadrados, isto é:

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \,\text{sen}(kx)\right) = S_n(x)\,,$$

tal que $Q = \min \|f(x) - S_n(x)\|^2$.

Para isso, vamos calcular a projeção ortogonal de $f(x)$ no subspaço gerado por:

$$\mathcal{B} = \left\{\frac{1}{2}, \cos(x), \,\text{sen}(x), \cos(2x), \,\text{sen}(2x), \ldots, \cos(nx), \,\text{sen}(nx)\right\}.$$

Adotando o produto interno $\langle f, g\rangle = \sum_{k=0}^{2N-1} f(x_k)g(x_k)$.

Aproximação trigonométrica: caso discreto

Teorema

O conjunto $\mathcal{B}$ é ortogonal, isto é:

$$\langle \operatorname{sen}(px), \cos(qx) \rangle = \sum_{k=0}^{2N-1} \operatorname{sen}(px_k)\cos(qx_k)dx = 0,\ \forall p, q \in \mathbb{N}$$

$$\langle \operatorname{sen}(px), \operatorname{sen}(qx) \rangle = \sum_{k=0}^{2N-1} \operatorname{sen}(px_k)\operatorname{sen}(qx_k)dx = \begin{cases} 0 & p \neq q \\ N & p = q \neq 0 \end{cases}$$

$$\langle \cos(px), \cos(qx) \rangle = \sum_{k=0}^{2N-1} \cos(px_k)\cos(qx_k)dx = \begin{cases} 0 & p \neq q \\ N & p = q \neq 0 \end{cases}$$

Aproximação trigonométrica: caso discreto

Teorema

O conjunto $\mathcal{B}$ é ortogonal, isto é:

$$\langle \operatorname{sen}(px), \cos(qx)\rangle = \sum_{k=0}^{2N-1} \operatorname{sen}(px_k)\cos(qx_k)dx = 0,\ \forall p, q \in \mathbb{N}$$

$$\langle \operatorname{sen}(px), \operatorname{sen}(qx)\rangle = \sum_{k=0}^{2N-1} \operatorname{sen}(px_k)\operatorname{sen}(qx_k)dx = \begin{cases} 0 & p \neq q \\ N & p = q \neq 0 \end{cases}$$

$$\langle \cos(px), \cos(qx)\rangle = \sum_{k=0}^{2N-1} \cos(px_k)\cos(qx_k)dx = \begin{cases} 0 & p \neq q \\ N & p = q \neq 0 \end{cases}$$

Demosntração: Livro *Análise Numérica*, Burden e Faires (seção 8.5).

Aproximação trigonométrica: caso discreto

Pelo teorema anterior, o sistema de equações normais é dado por:

Aproximação trigonométrica: caso discreto

Pelo teorema anterior, o sistema de equações normais é dado por:

$$\begin{bmatrix} \frac{N}{2} & 0 & \cdots & 0 \\ 0 & N & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \left\langle \frac{1}{2}, f \right\rangle \\ \langle \cos(x), f \rangle \\ \vdots \\ \langle \operatorname{sen}(nx), f \rangle \end{bmatrix}$$

Aproximação trigonométrica: caso discreto

Pelo teorema anterior, o sistema de equações normais é dado por:

$$\begin{bmatrix} \frac{N}{2} & 0 & \cdots & 0 \\ 0 & N & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \langle \frac{1}{2}, f \rangle \\ \langle \cos(x), f \rangle \\ \vdots \\ \langle \operatorname{sen}(nx), f \rangle \end{bmatrix}$$

Portanto,

$$\begin{aligned} a_0 &= \frac{1}{N} \sum_{k=0}^{2N-1} f(x_k) \\ a_j &= \frac{1}{N} \sum_{k=0}^{2N-1} f(x_k) \cos(jx_k), \quad j = 1, \ldots, n \\ b_j &= \frac{1}{N} \sum_{k=0}^{2N-1} f(x_k) \operatorname{sen}(jx_k), \quad j = 1, \ldots, n \end{aligned}$$

Aproximação trigonométrica: caso discreto

Exemplo

Seja $f(x)$ uma função discreta dada pela tabela abaixo:

x	$-\pi$	$-2\pi/3$	$-\pi/3$	0	$\pi/3$	$-2\pi/3$
$f(x)$	10.74	-0.23	-6.81	-9.00	-6.81	-0.23

Gostaríamos de aproximar $f(x)$ por $S_1(x)$ no sentido dos mínimos quadrados, ou seja:

$$f(x) \approx \frac{a_0}{2} + a_1 \cos(x) + b_1 \operatorname{sen}(x) = S_1(x)$$

Aproximação trigonométrica: caso discreto

Exemplo (continuação)

Sabendo que $2N = 6$, logo os coeficientes de $S_1(x)$ são:

Aproximação trigonométrica: caso discreto

Exemplo (continuação)

Sabendo que $2N = 6$, logo os coeficientes de $S_1(x)$ são:

$$a_0 = \frac{1}{3}\sum_{k=0}^{5} f(x_k) = -4.11$$

Aproximação trigonométrica: caso discreto

Exemplo (continuação)

Sabendo que $2N = 6$, logo os coeficientes de $S_1(x)$ são:

$$
\begin{aligned}
a_0 &= \frac{1}{3}\sum_{k=0}^{5} f(x_k) = -4.11 \\
a_1 &= \frac{1}{3}\sum_{k=0}^{5} f(x_k)\cos(x_k) = -8.77
\end{aligned}
$$

Aproximação trigonométrica: caso discreto

Exemplo (continuação)

Sabendo que $2N = 6$, logo os coeficientes de $S_1(x)$ são:

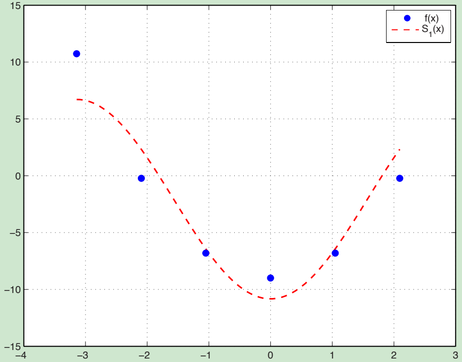
$$
\begin{aligned}
a_0 &= \frac{1}{3}\sum_{k=0}^{5} f(x_k) = -4.11 \\
a_1 &= \frac{1}{3}\sum_{k=0}^{5} f(x_k)\cos(x_k) = -8.77 \\
b_1 &= \frac{1}{3}\sum_{k=0}^{5} f(x_k)\,\mathrm{sen}(x_k) \approx 0
\end{aligned}
$$

Portanto,

$$f(x) \approx -2.06 - 8.77\cos(x)$$

Aproximação trigonométrica: caso discreto

Exemplo (continuação)



Aproximação de $f(x)$ por $S_1(x) = -2.06 - 8.77\cos(x)$.

Mínimos quadrados móveis (MLS)

Vamos considerar f uma função discreta definida em $n+1$ pontos.
Seja V um espaço vetorial gerado por $\{\varphi_0, \ldots, \varphi_m\}$, com $m < n$.

Mínimos quadrados móveis (MLS)

Vamos considerar f uma função discreta definida em $n+1$ pontos. Seja V um espaço vetorial gerado por $\{\varphi_0, \ldots, \varphi_m\}$, com $m < n$.

Novamente, queremos aproximar f por uma função $F^* \in V$ através de mínimos quadrados ponderados.

Mínimos quadrados móveis (MLS)

Vamos considerar f uma função discreta definida em $n+1$ pontos. Seja V um espaço vetorial gerado por $\{\varphi_0, \ldots, \varphi_m\}$, com $m < n$.

Novamente, queremos aproximar f por uma função $F^* \in V$ através de mínimos quadrados ponderados.

Para isso, vamos definir uma *função peso* através da função gaussiana:

$$w(x) = e^{-\frac{(x-c)^2}{\sigma^2}} = \psi\left(\frac{|x-c|}{\sigma}\right), \quad \text{com} \quad \psi(t) = e^{-t^2}$$

onde c é o *centro* da gaussiana e o raio é controlado por σ.

Mínimos quadrados móveis (MLS)

Vamos considerar f uma função discreta definida em $n+1$ pontos. Seja V um espaço vetorial gerado por $\{\varphi_0, \ldots, \varphi_m\}$, com $m < n$.

Novamente, queremos aproximar f por uma função $F^* \in V$ através de mínimos quadrados ponderados.

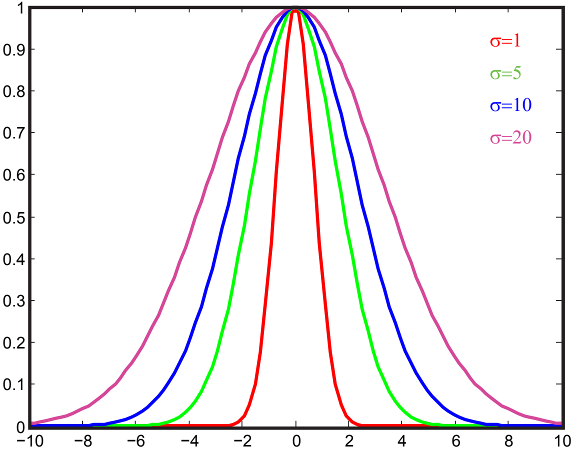
Para isso, vamos definir uma *função peso* através da função gaussiana:

$$w(x) = e^{-\frac{(x-c)^2}{\sigma^2}} = \psi\left(\frac{|x-c|}{\sigma}\right), \quad \text{com} \quad \psi(t) = e^{-t^2}$$

onde c é o *centro* da gaussiana e o raio é controlado por σ.

Quando movemos o centro da gaussiana, privilegiamos os pontos mais próximos a c, com vizinhança controlada por σ.

Função peso



Mínimos quadrados móveis (MLS)

Assim, temos que $\mathbf{w}(x) = (w_0(x), \ldots, w_n(x))^\top$ com

$$w_i(x) = \psi\left(\frac{|x - x_i|}{\sigma}\right).$$

Mínimos quadrados móveis (MLS)

Assim, temos que $\mathbf{w}(x) = (w_0(x), \ldots, w_n(x))^\top$ com

$$w_i(x) = \psi\left(\frac{|x - x_i|}{\sigma}\right).$$

Podemos definir o produto interno:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}(x)} = \sum_{i=0}^{n} w_i(x) u_i v_i$$

Mínimos quadrados móveis (MLS)

Assim, temos que $\mathbf{w}(x) = (w_0(x), \ldots, w_n(x))^\top$ com

$$w_i(x) = \psi\left(\frac{|x - x_i|}{\sigma}\right).$$

Podemos definir o produto interno:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}(x)} = \sum_{i=0}^{n} w_i(x) u_i v_i$$

Logo, a melhor aproximação para f no sentido dos mínimos quadrados é a função $F^* = \alpha_0^* \varphi_0 + \cdots + \alpha_m^* \varphi_m$, cujos coeficientes são solução do sistema linear $\mathbf{A}(x)\alpha^*(x) = \mathbf{b}(x)$,

Mínimos quadrados móveis (MLS)

Assim, temos que $\mathbf{w}(x) = (w_0(x), \ldots, w_n(x))^\top$ com

$$w_i(x) = \psi\left(\frac{|x - x_i|}{\sigma}\right).$$

Podemos definir o produto interno:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}(x)} = \sum_{i=0}^{n} w_i(x) u_i v_i$$

Logo, a melhor aproximação para f no sentido dos mínimos quadrados é a função $F^* = \alpha_0^* \varphi_0 + \cdots + \alpha_m^* \varphi_m$, cujos coeficientes são solução do sistema linear $\mathbf{A}(x)\alpha^*(x) = \mathbf{b}(x)$, onde:

$$\mathbf{A}(x) = \begin{bmatrix} \langle \varphi_0, \varphi_0 \rangle_{\mathbf{w}(x)} & \cdots & \langle \varphi_0, \varphi_m \rangle_{\mathbf{w}(x)} \\ \vdots & & \vdots \\ \langle \varphi_m, \varphi_0 \rangle_{\mathbf{w}(x)} & \cdots & \langle \varphi_m, \varphi_m \rangle_{\mathbf{w}(x)} \end{bmatrix} \text{ e } \mathbf{b}(x) = \begin{bmatrix} \langle f, \varphi_0 \rangle_{\mathbf{w}(x)} \\ \vdots \\ \langle f, \varphi_m \rangle_{\mathbf{w}(x)} \end{bmatrix}$$

Aproximação polinomial

No caso em que $\{\varphi_0 = 1, \varphi_1 = x, \ldots, \varphi_m = x^m\}$, podemos escrever o sistema linear $\mathbf{A}(x)\alpha^*(x) = \mathbf{b}(x)$ da seguinte forma:

Aproximação polinomial

No caso em que $\{\varphi_0 = 1, \varphi_1 = x, \ldots, \varphi_m = x^m\}$, podemos escrever o sistema linear $\mathbf{A}(x)\alpha^*(x) = \mathbf{b}(x)$ da seguinte forma:

$$\mathbf{M}(x)\alpha^*(x) = \mathbf{X}^\top \mathbf{W}(x)\mathbf{y}$$

com

$$\mathbf{M}(x) = \mathbf{X}^\top \mathbf{W}(x)\mathbf{X} \quad \text{e} \quad \mathbf{W}(x) = \mathrm{diag}(w_0(x), w_1(x), \ldots, w_n(x))$$

Aproximação polinomial

No caso em que $\{\varphi_0 = 1, \varphi_1 = x, \ldots, \varphi_m = x^m\}$, podemos escrever o sistema linear $\mathbf{A}(x)\alpha^*(x) = \mathbf{b}(x)$ da seguinte forma:

$$\mathbf{M}(x)\alpha^*(x) = \mathbf{X}^\top \mathbf{W}(x)\mathbf{y}$$

com

$$\mathbf{M}(x) = \mathbf{X}^\top \mathbf{W}(x)\mathbf{X} \quad \text{e} \quad \mathbf{W}(x) = \text{diag}(w_0(x), w_1(x), \ldots, w_n(x))$$

Portanto, a solução do sistema é dada por:

$$\alpha^*(x) = \mathbf{M}(x)^{-1}\mathbf{X}^\top \mathbf{W}(x)\mathbf{y}$$