Scuola Interdipartimentale delle **Scienze**, dell'**Ingegneria** e della **Salute**

L. Magistrale in IA (ML&BD)

Scientific Computing
(part 2 – 6 credits)

prof. Mariarosaria Rizzardi

Centro Direzionale di Napoli – Bldg. C4
room: n. 423 – North Side, 4th floor
phone: 081 547 6545
email: mariarosaria.rizzardi@uniparthenope.it

Contents

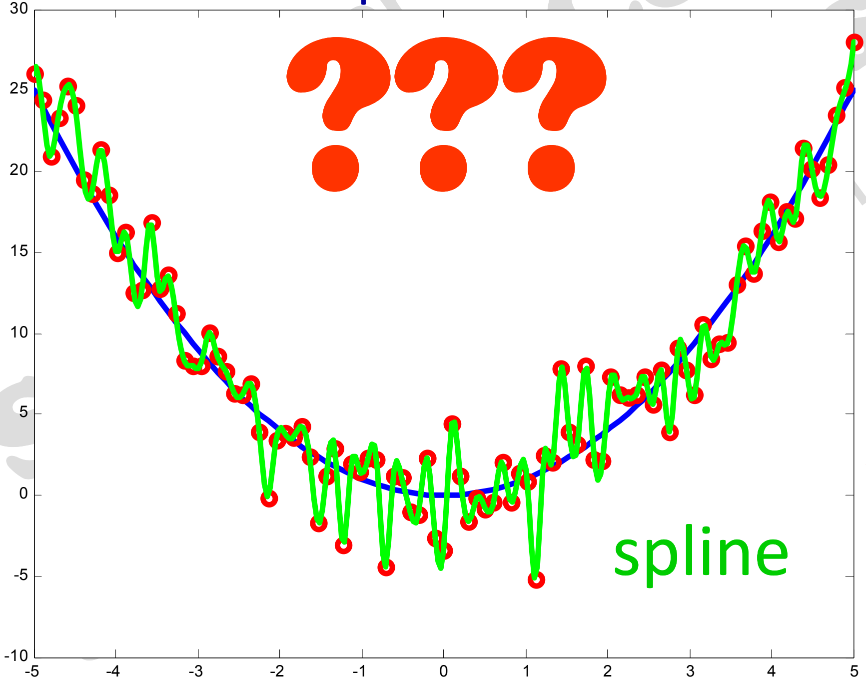
- Interpolation VS approximation.
- Best approximation in $\|\cdot\|_2$: finite dimension subspaces.

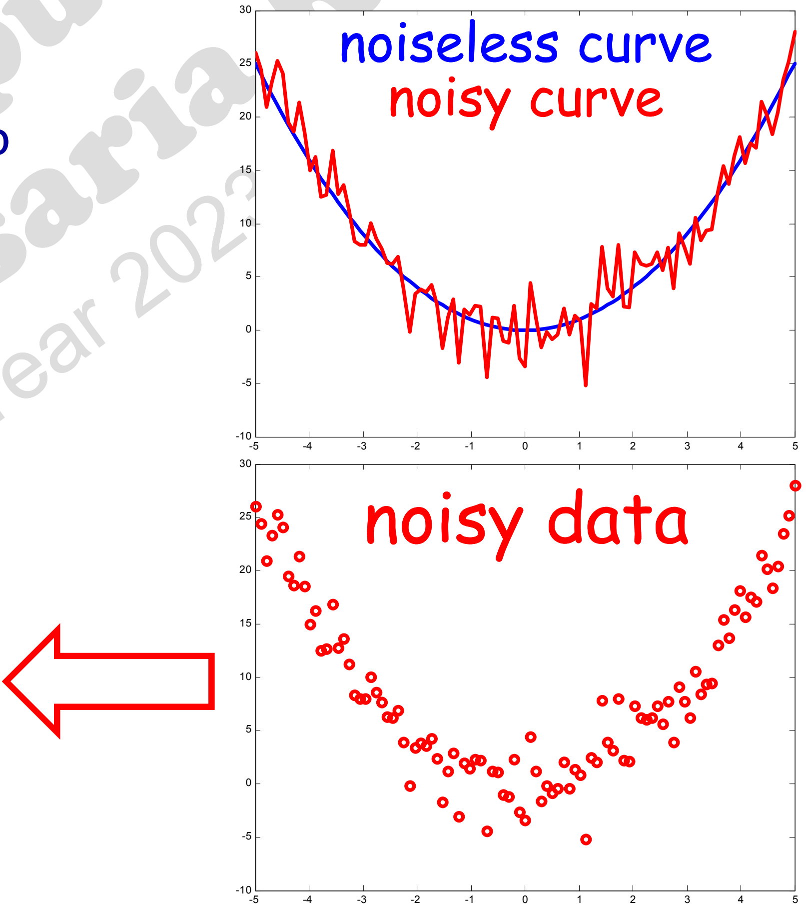
Recall

Data interpolation

A function interpolating the values of an unknown function, in general, is not a good approximation of the function that generated the data.

If data are perturbed, Interpolation also reconstructs the perturbations!

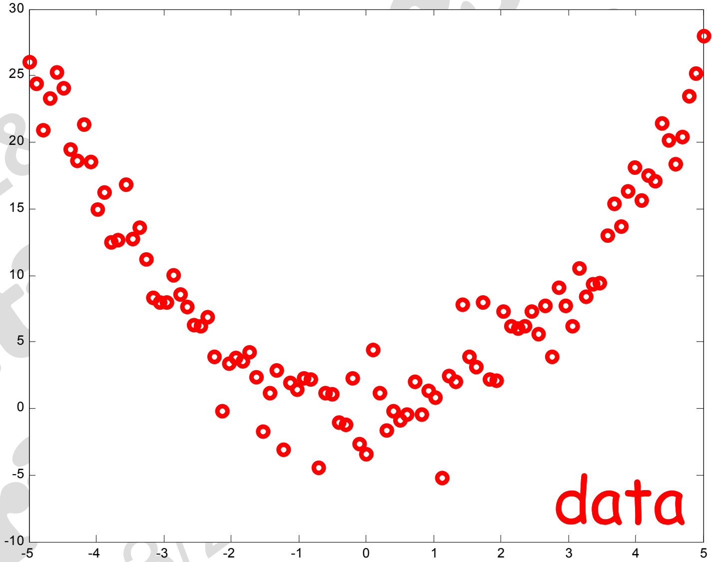




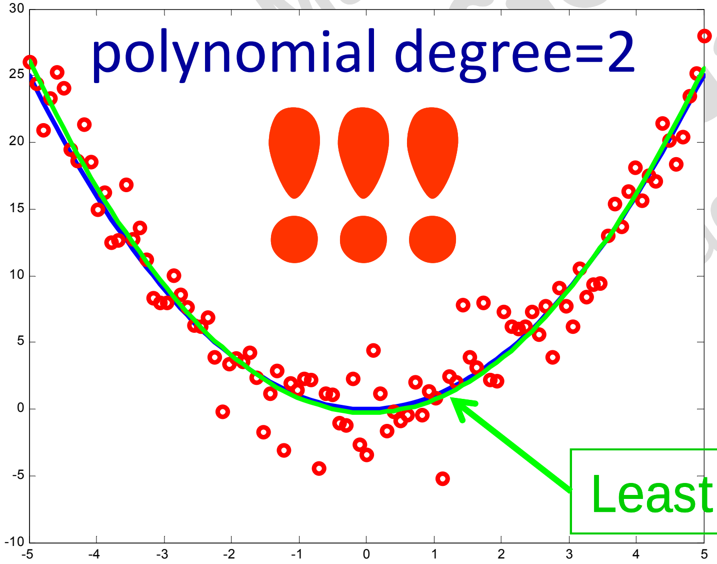
Recall

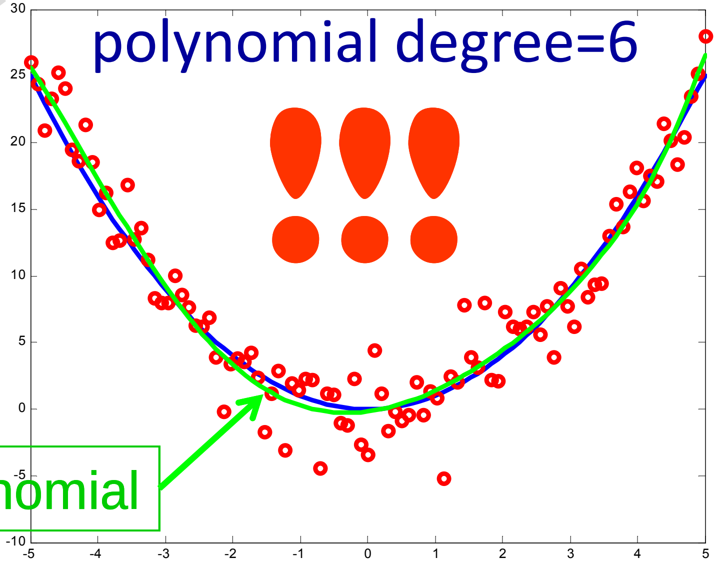
Data fitting

If data are perturbated, **Best Fit** reduces perturbations!



Fitting by means of a Least Squares polynomial





Least Squares polynomial

Best approximation in normed Linear Spaces

Let $\mathbf{X}$ be a Linear Space, M a subspace, $M\subset\mathbf{X}$, and $f\in\mathbf{X}$, but $f\notin M$. A vector $f^*\in M$ is said the **best approximation** of f w.r.t. $\|\cdot\|$:

$$\overset{\text{def}}{\Longleftrightarrow} \qquad \|f-f^*\| = \min_{g\in M}\|f-g\|$$

$$f^* = \arg\min_{g\in M}\|f-g\|$$

that is, f^* is the “closest” approximation on M to f w.r.t. the selected norm.

1. discrete and finite case (vector $x \in \mathbb{R}^n$)

Euclidean or 2-norm

$$\|x\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}$$

uniform or ∞-norm

$$\|x\|_\infty = \max_{1 \le k \le n} |x_k|$$

Taxicab or 1-norm
or Manhattan norm

$$\|x\|_1 = \sum_{k=1}^{n} |x_k|$$

2. discrete and infinite case (sequence $\{x_k\}_k \in \ell^p$)

p-power summable sequences $\ell^p = \left\{ \{x_k\}_k : \boxed{\sum_k |x_k|^p < \infty} \right\}$

convergent series

p-norm $\|x\|_p = \sqrt[p]{\sum_{k=1}^{\infty} |x_k|^p}$

For $p=2$ we obtain again the Euclidean norm defined in ℓ^2, the Linear Space of square-summable sequences. In ℓ^2 the *standard scalar product*, which induces the Euclidean norm, is $\langle x, y \rangle = \sum_{k=1}^{\infty} \bar{x}_k y_k \quad \forall x_k, y_k \in \mathbb{C}$

3. continuous case

(function $f \in \ldots$)

2-norm $\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 \, dx}$ $f \in L^2[a,b]$

square-summable in $[a,b]$

∞-norm $\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$ $f \in C^0[a,b]$

continuous in $[a,b]$

1-norm $\|f\|_1 = \int_a^b |f(x)| \, dx$ $f \in L^1[a,b]$

summable in $[a,b]$

In the continuous case, the Euclidean norm is defined in the Hilbert Space $L^2[a,b]$, that is the space of square-integrable (or square-summable) functions; the norm is induced by the scalar product defined as

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

The integral has to be intended as Lebesgue integral; $L^2[a,b]$ contains equivalence classes of functions (two functions that coincide "almost everywhere", that is except on a set of zero measure according to Lebesgue measure, are considered the same); $L^2[a,b]$ is **complete**, while $C^2[a,b]$ does not.

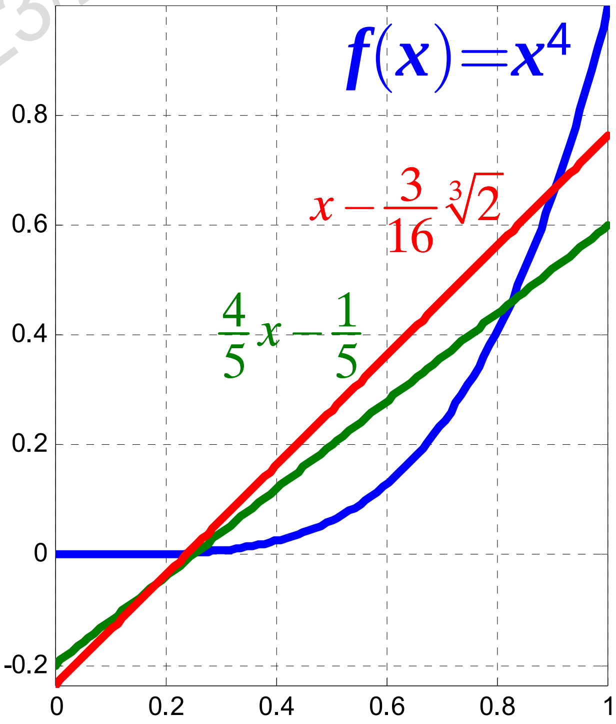
The best approximation depends on:

- the subspace where it is searched (… of course!)
- the selected norm

Example

The best approximation of the function $f(x)=x^4$, $x\in[0,1]$, in the subspace of first degree polynomials is given by:

- $\frac{4}{5}x-\frac{1}{5}$ w.r.t. $\|\cdot\|_2$
- $x-\frac{3}{16}\sqrt[3]{2}$ w.r.t. $\|\cdot\|_\infty$



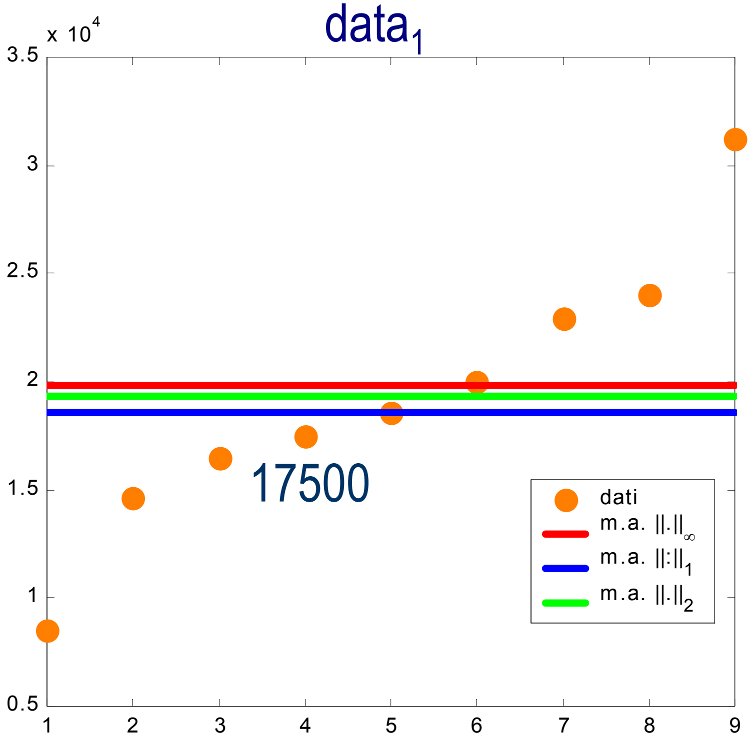
In the case of discrete data the best approximation also depends on:

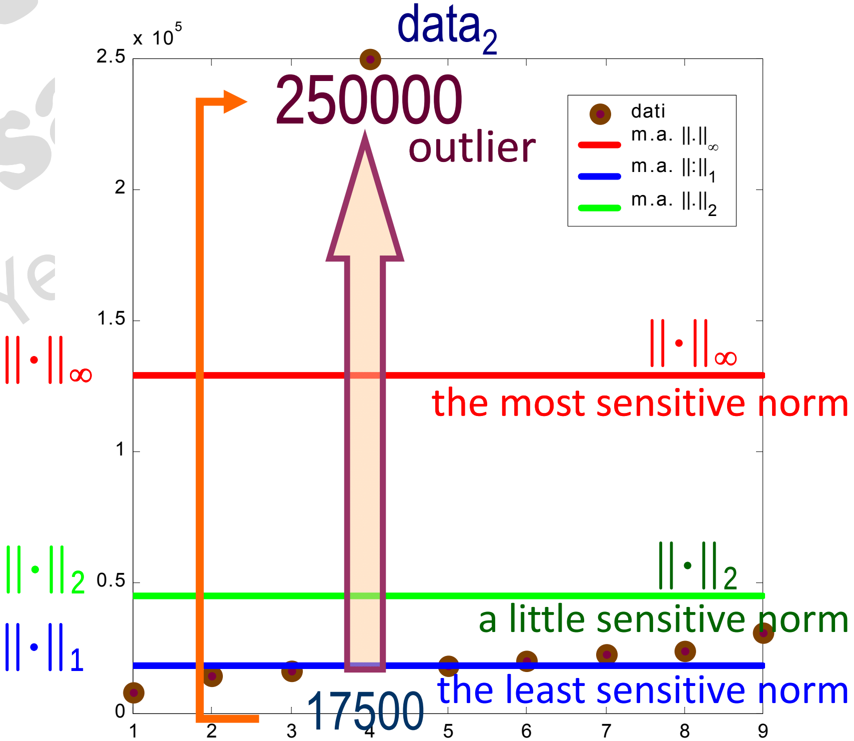
- the particular data

Example

All the samples must be approximated by a single value

$data_1$ = (8500,14600,16500, 17500, 18600,20000,22900,24000,31200)
$data_2$ = (8500,14600,16500, 250000, 18600,20000,22900,24000,31200)





In $data_2$ a single value has been modified very much ...

Linear least squares (LLS) approximation

(best linear approximation in $\|\cdot\|_2$)

simpler to compute numerically

Let: $\mathbf{X}$ be a linear space eqipped with $\|\cdot\|_2$ induced by $\langle\cdot,\cdot\rangle$,
f be a function $f\in\mathbf{X}$,
M_n a subspace of $\mathbf{X}$ with finite dimension ($\dim M_n = n<\infty$),
whose basis is known (M_n= span{$\varphi_1, \varphi_2, \ldots, \varphi_n$}).

Theorem for existence and uniqueness

of the best approximation w.r.t. $\|\cdot\|_2$ in a finite dimension subspace

The problem of the **best linear approximation in $\|\cdot\|_2$ of f** in M_n ($\dim M_n<\infty$) admits **only one solution** $f_n^*\in M_n \Rightarrow f_n^* = \sum_{k=1}^{n} c_k^*\varphi_k$

such that

$$f_n^* = \arg\min_{g_n\in M_n}\|f-g_n\|_2 \Leftrightarrow \|f-f_n^*\|_2 = \min_{g_n\in M_n}\|f-g_n\|_2$$

if, and only if,

$$\langle f-f_n^*, g_n\rangle = 0 \qquad \forall g_n\in M_n$$

(i.e., if the residual vector $f-f_n^*$ is **orthogonal to the subspace M_n**).

(intuitive) Geometrical interpretation of the best approximation in $\|\cdot\|_2$: orthogonal projection

(previous Theorem)

Theorem for existence and uniqueness

(Theor. of Normal Equations)

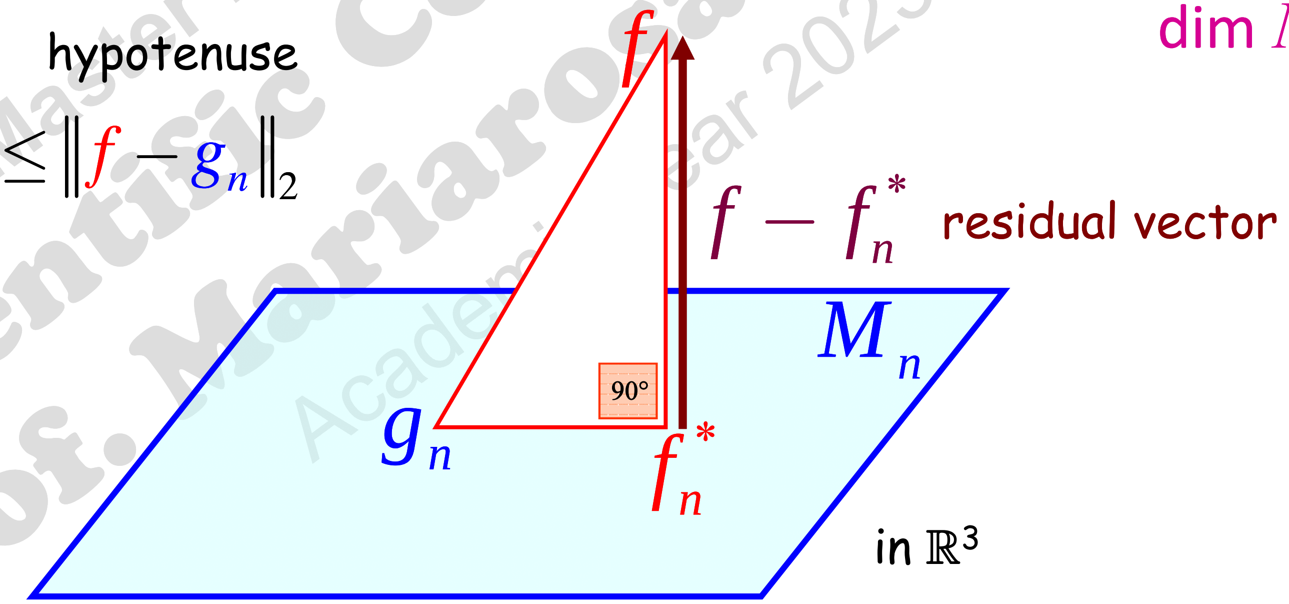
f_n^* = the best linear approximation (w.r.t. $\|\cdot\|_2$) of f in M_n

$$\Longleftrightarrow \quad \langle f - f_n^*, g_n \rangle = 0 \qquad \forall g_n \in M_n$$

dim $M_n < \infty$

cathetus hypotenuse

$$\|f - f_n^*\|_2 \leq \|f - g_n\|_2$$



The best approximation f_n^* of f in M_n is the orthogonal projection of f onto M_n

The orthogonality condition to the subspace $\boldsymbol{M}_n$, of finite dimension n and such that $\boldsymbol{M}_n = \mathrm{span}\{\varphi_1, \varphi_2, \dots, \varphi_n\}$, is equivalent to the linear system of n equations with unknowns c_k^*:

element n. i of the basis

Theor. of Normal Eqs

$$\langle f - f_n^*, \varphi_i\rangle = 0 \qquad i=1,2,\dots,n$$

⇨ If we put $f_n^* = \sum_{k=1}^{n} c_k^* \varphi_k$ in the previous $\langle\cdot,\cdot\rangle$ and reorganize the equations, we get

component of f along φ_i

$$\boxed{\sum_{k=1}^{n} \boxed{\langle \varphi_i, \varphi_k\rangle} c_k^* = \langle f, \varphi_i\rangle \qquad i = 1, 2, \dots, n}$$

Gram matrix

⇧

Linear System of Normal Equations

Best linear approximation in $\|\cdot\|_2$

➢finite dimension subspace

Theor. ⇒
- discrete case
- continuous case

The previous Theorem ensures that the solution exists and is unique

➢infinite dimension subspace

no Theor. $\cancel{\Rightarrow}$
- discrete case
- continuous case

?

WE'LL SEE LATER

Best linear approximation in $\|\cdot\|_2$

finite dimension:

discrete case

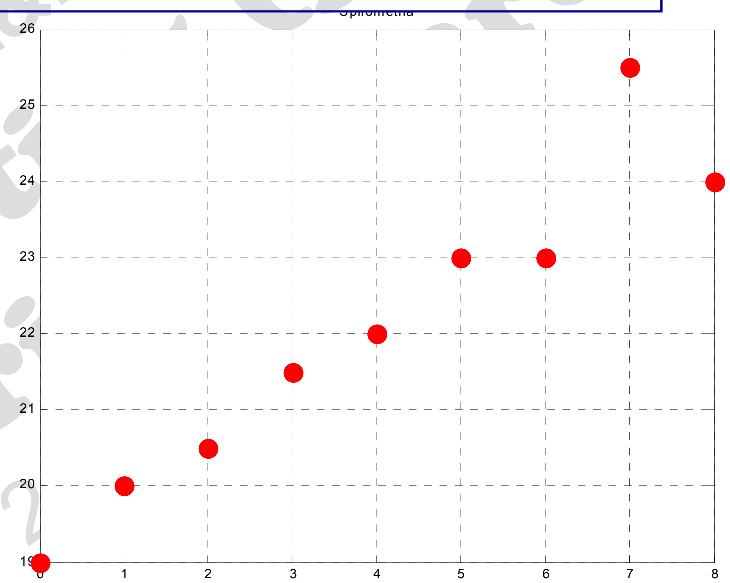
incompatible systems
overdetermined systems

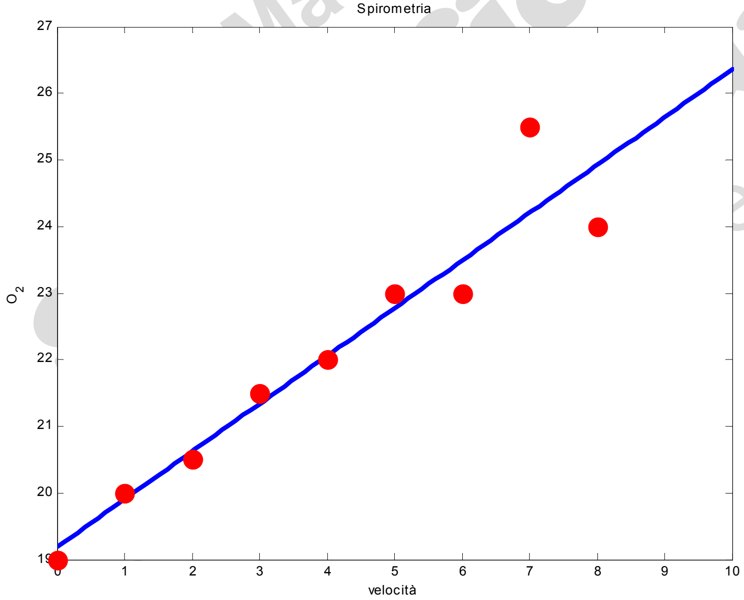
Discrete case

space $\mathbb{R}^n$ with $\langle x,y \rangle = x^\mathsf{T} y$

Example of application in $\mathbb{R}^2$

Spirometry measures the oxygen diffusion capacity in the lungs. The graph shows some results as a function of the speed at which the patient is moving.



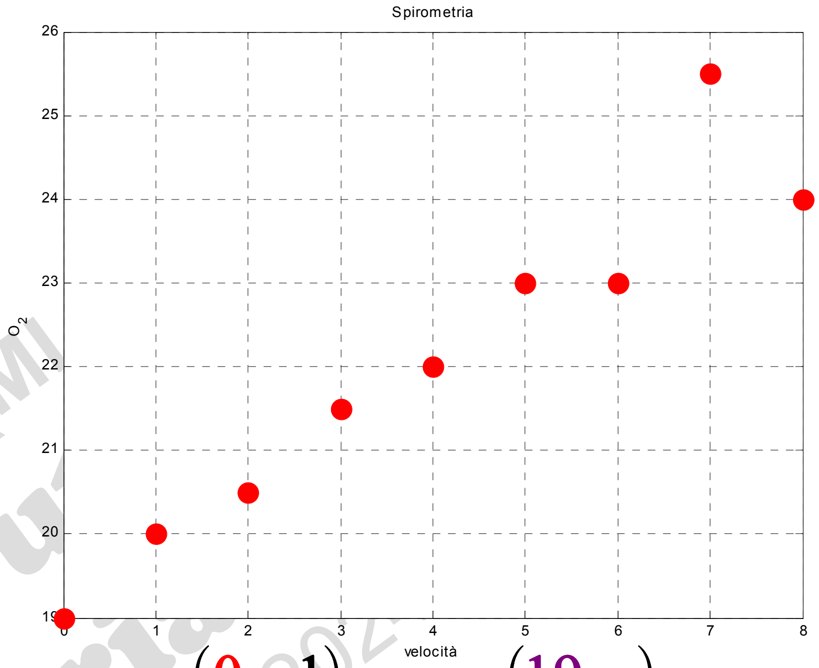


The mathematical model for the dependence between oxygen and velocity is linear

$$f(x) = mx + q$$

however ... samples are not aligned!!!

Data points are not aligned!



There is no line ***r***

$$r : y=mx+q$$

passing through all the points, i.e. such that

$$m\,x_i + q = y_i \qquad (i=1,2,\ldots n)$$

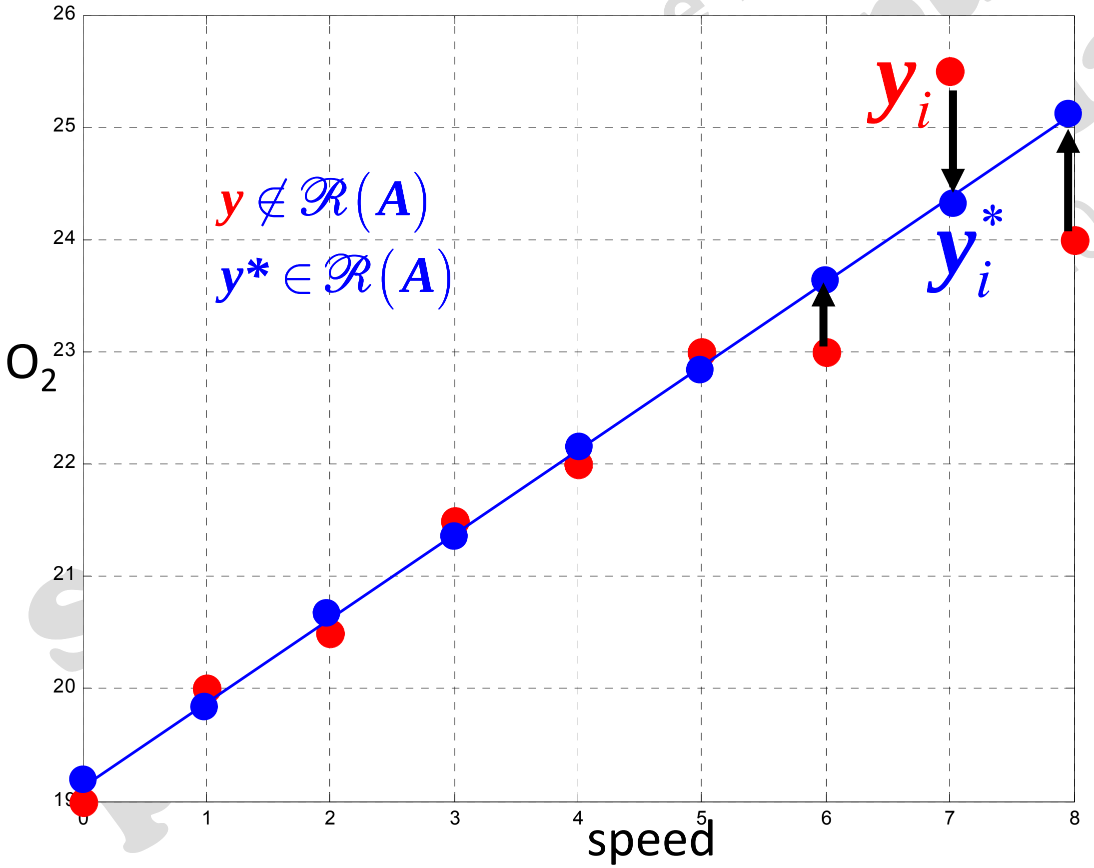
$$\underset{A}{\begin{pmatrix}0 & 1\\ 1 & 1\\ 2 & 1\\ 3 & 1\\ 4 & 1\\ 5 & 1\\ 6 & 1\\ 7 & 1\\ 8 & 1\end{pmatrix}}\underset{u}{\begin{pmatrix}m\\ q\end{pmatrix}}=\underset{y}{\begin{pmatrix}19\\ 20\\ 20.5\\ 21.5\\ 22\\ 23\\ 23\\ 25.5\\ 24\end{pmatrix}}$$

$y \notin \mathcal{R}(A)$

$Au=y$ is an incompatible linear system because of data errors

How to find the line?

The polynomial Least Squares Method (LSM) replaces the known term $\boldsymbol{y}$ of the system $\boldsymbol{Au}=\boldsymbol{y}$ with another vector $\boldsymbol{y}^*$ which makes the system **compatible**. LSM approximates $\boldsymbol{y}$ by $\boldsymbol{y}^*$, where $\boldsymbol{y}^*$ is the closest vector on $\mathscr{R}(\boldsymbol{A})$ to $\boldsymbol{y}$, i.e. $\boldsymbol{y}^*$ is the best approximation of $\boldsymbol{y}$ in $\mathscr{R}(\boldsymbol{A})$.



$$\mathscr{R}(\boldsymbol{A}) = \mathrm{span}\left\{\begin{pmatrix}0\\1\\2\\3\\4\\5\\6\\7\\8\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\\1\\1\\1\\1\\1\end{pmatrix}\right\}$$

System of Normal Equations

$$\boxed{A^{\mathsf{T}} A\, u^* = A^{\mathsf{T}} y}$$

always compatible

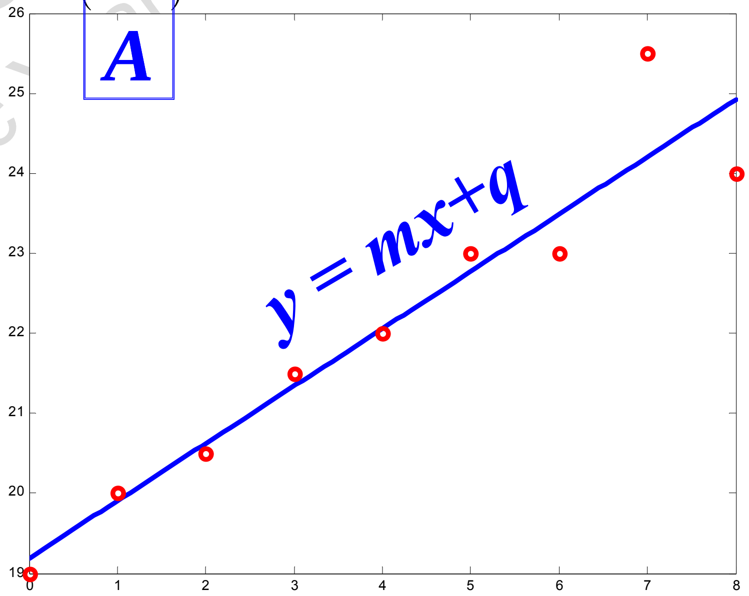
why?

$$\sum_{k=1}^{n} \langle \varphi_i, \varphi_k \rangle c_k^* = \langle f, \varphi_i \rangle \qquad i = 1, 2, \dots, n$$

$$\underbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}}_{A^{\mathsf{T}}} \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \end{pmatrix}}_{A} \begin{pmatrix} m \\ q \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 19 \\ 20 \\ 20.5 \\ 21.5 \\ 22 \\ 23 \\ 23 \\ 25.5 \\ 24 \end{pmatrix}$$

$$\begin{pmatrix} 204 & 36 \\ 36 & 9 \end{pmatrix} \begin{pmatrix} m \\ q \end{pmatrix} = \begin{pmatrix} 837 \\ 198.5 \end{pmatrix}$$

$$\begin{pmatrix} m \\ q \end{pmatrix} = \begin{pmatrix} 0.7167 \\ 19.1889 \end{pmatrix}$$



... in *MATLAB*

```
xi=(0:8)'; A=[xi  ones(9,1)]; yi=[19 20 20.5 21.5 22 23 23 25.5 24]';
disp([rank(A) rank([A yi])])
2     3    incompatible system  A*c=yi
```

```
c = A \ yi
c =
    0.7167
   19.1889
m=c(1); q=c(2); y_star=m*xi+q;
plot(xi,yi,'or',xi,y_star,'ob')
```

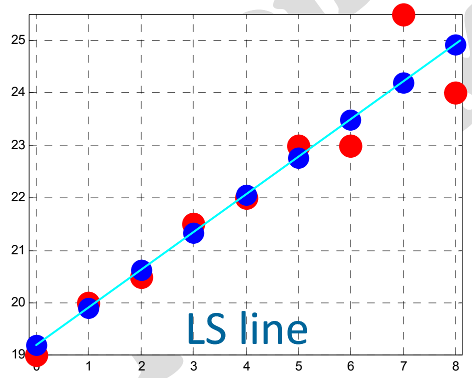
```
c = polyfit(xi,yi,1)
c =
    0.7167
   19.1889
y_star = polyval(c,xi);
plot(xi,yi,'or',xi,y_star,'ob')
```

If the linear system **A*c=yi** is incompatible and rank(A)=n (with n=number of columns in A, n<m), to solve the system, the statement **c=A\yi** returns the only least squares solution.

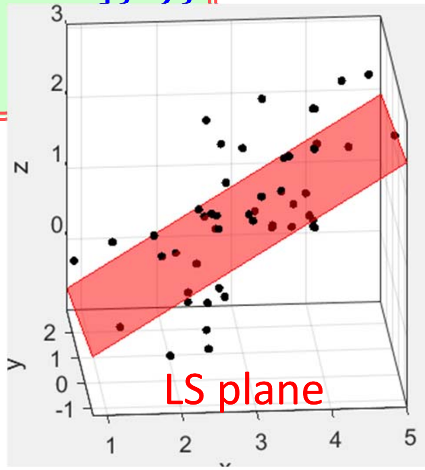
Only for $\mathbb{R}^2$, the **polyfit()** function returns the coefficients of the best polynomial approximation (in a least-squares sense) for the data in **yi**.

Random Multivariate Normal Distribution

```
N=50; P=mvnrnd([3 1 1],[1 .2 .7;.2 1 0;.7 0 1],N);
Xi=P(:,1); Yi=P(:,2); Zi=P(:,3);
A=[Xi Yi ones(N,1)];   c = A \ Zi;
```



In this case ($\mathbb{R}^3$), we cannot use **polyfit** function.



General discrete case

Resolution of a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad \mathbf{A}_{(m\times n)}$$

by means of **Linear Least Squares method**, where

- A is a rectangular matrix, with rank n, and $n < m$
- The system is incompatible

The **Least Squares Method** replaces the incompatible system $\boldsymbol{Ax=b}$, $\boldsymbol{b} \notin \mathscr{R}(\boldsymbol{A})$, with a compatible system

$$\boldsymbol{Ax} = \boldsymbol{p}^* \quad (\boldsymbol{p}^* \in \mathscr{R}(\boldsymbol{A}))$$

where $\boldsymbol{p}^* = \sum_{k=1}^{n} c_k^* \boldsymbol{A}_{.,k}$ is the **best approximation** of $\boldsymbol{b}$ in $\mathscr{R}(\boldsymbol{A})$, i.e. it is the vector of $\mathscr{R}(\boldsymbol{A})$ which is "closest to $\boldsymbol{b}$" w.r.t. $\|\cdot\|_2$.

DEF $\boldsymbol{x}^*$ = Least Squares solution of the incompatible system $\boldsymbol{Ax=b}$ $\overset{def}{\Longleftrightarrow}$

$$\boldsymbol{x}^* : \boldsymbol{Ax}^* = \boldsymbol{p}^* \qquad (\boldsymbol{x}^* \text{ solves the system } \boldsymbol{Ax} = \boldsymbol{p}^*)$$

where

$$\boldsymbol{p}^* = \arg\min_{p \in \mathscr{R}(\boldsymbol{A})} \|\boldsymbol{b} - \boldsymbol{p}\|_2$$

… and if the system is compatible?

incompatible system

```
A=[1 2; 1 5; 0 0];
b=[4 3 9]';
disp([rank(A) rank([A b])])
2     3
c=A\b;    d=(A'*A)\(A'*b);
disp([c d])
    4.6667    4.6667
   -0.3333   -0.3333
```

compatible system

```
A=[1 2; 1 5; 0 0];
b=[4 3 0]';
disp([rank(A) rank([A b])])
2     2
c=A\b;      d=(A'*A)\(A'*b);
disp([c d])
    4.6667    4.6667
   -0.3333   -0.3333
```

underdetermined compatible system

```
A=[1 2; 1 2; 0 0];
b=[4 4 0]';
disp([rank(A) rank([A b])])
1     1
c=A\b;      d=(A'*A)\(A'*b);
disp([c d])
         0    0.0000
    2.0000    2.0000
```

In this case, the Least Squares method returns

the only solution,

or

one of the ∞ solutions of the system.

Algorithm

In order to solve an incompatible system $Ax=b$ by means of Least Squares method ...

1. We solve the System of Normal Equations

$$A^T A c^* = A^T b$$

2. We compute $p^* = \sum_{k=1}^{n} c_k^* A_{\cdot,k}$ $\Longleftrightarrow$ $\boxed{Ac^* = p^*}$

3. Then, to get the Least Squares solution x^*, we solve the compatible system $Ax^* = p^*$.

but, ... by definition of Least Squares solution ...

Property

The Least Squares solution x^* of the incompatible system $Ax=b$ is the same as the solution c^* of Normal Equations.

In the previous algorithm, among the three steps, only the step 1. is required; that is, in order to compute the LS solution x^*, it suffices to solve the system of Normal Equations.

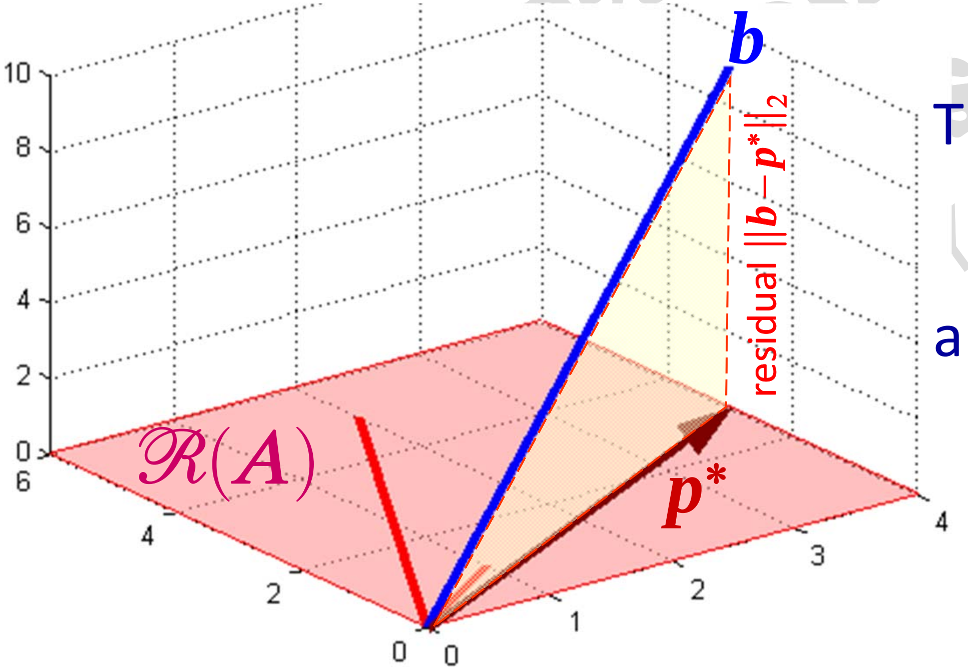
Example 1

$$\begin{cases} x+2y=4 \\ x+5y=3 \\ 0x+0y=9 \end{cases} \Longleftrightarrow \quad \boldsymbol{Ax=b} \qquad \boldsymbol{A}=\begin{pmatrix}1&2\\1&5\\0&0\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}4\\3\\9\end{pmatrix}$$

incompatible

$$\boldsymbol{b}\notin\mathscr{R}(\boldsymbol{A})=\text{span}\{(1,1,0)^\mathsf{T},(2,5,0)^\mathsf{T}\}$$

$$\boldsymbol{p}^* \;:\; \left\|\boldsymbol{b}-\boldsymbol{p}^*\right\|_2 = \min_{p\in\mathscr{R}(\boldsymbol{A})}\left\|\boldsymbol{b}-\boldsymbol{p}\right\|_2$$



The Normal Equations

$$\sum_{k=1}^{n}\langle\varphi_i,\varphi_k\rangle c_k^* = \langle f,\varphi_i\rangle \quad i=1,2,\dots,n$$

are written in matrix form as

$$\boxed{\boldsymbol{A}^\mathsf{T}\boldsymbol{A}}\,\boldsymbol{x}^* = \boldsymbol{A}^\mathsf{T}\boldsymbol{b}$$

Gram matrix

$\boldsymbol{p}^*$ is the orthogonal projection of $\boldsymbol{b}$ onto $\mathscr{R}(\boldsymbol{A})$

Contents

- **Best linear approximation in finite dimension subspaces w.r.t. $\|\cdot\|_2$: continuous case.**

Example 2

Compute the best linear approximation $f^*(x)$ of $f(x)=x^3$ w.r.t. $\|\cdot\|_2$ in the subspace $\Pi_1[-1,+1]$ of 1st degree algebraic polynomials over $[-1,+1]$: $M_n = \Pi_1[-1,+1] = \mathrm{span}\{1, x\}$

Space $C[-1,+1]$ with $\langle f,g\rangle = \int_{-1}^{+1} f(x)g(x)\,\mathrm{d}x$

$$f^* \in \Pi_1[-1,1] \qquad f^*(x) = c_1^* + c_2^* x$$

unknowns

$$\begin{cases} \langle \varphi_1,\varphi_1\rangle c_1^* + \langle \varphi_1,\varphi_2\rangle c_2^* = \langle f,\varphi_1\rangle \\ \langle \varphi_2,\varphi_1\rangle c_1^* + \langle \varphi_2,\varphi_2\rangle c_2^* = \langle f,\varphi_2\rangle \end{cases}$$

Normal Equations

$$\begin{cases} \left[\int_{-1}^{1} 1\,\mathrm{d}x\right] c_1^* + \left[\int_{-1}^{1} 1\boldsymbol{x}\,\mathrm{d}x\right] c_2^* = \left[\int_{-1}^{1} \boldsymbol{x}^3 1\,\mathrm{d}x\right] \\ \left[\int_{-1}^{1} 1\boldsymbol{x}\,\mathrm{d}x\right] c_1^* + \left[\int_{-1}^{1} \boldsymbol{x}^2\,\mathrm{d}x\right] c_2^* = \left[\int_{-1}^{1} \boldsymbol{x}^3 \boldsymbol{x}\,\mathrm{d}x\right] \end{cases}$$

$$\begin{cases} c_1^* = 0 \\ c_2^* = \dfrac{3}{5} \end{cases}$$

$$\boxed{f^*(\boldsymbol{x}) = \frac{3}{5}\boldsymbol{x}}$$

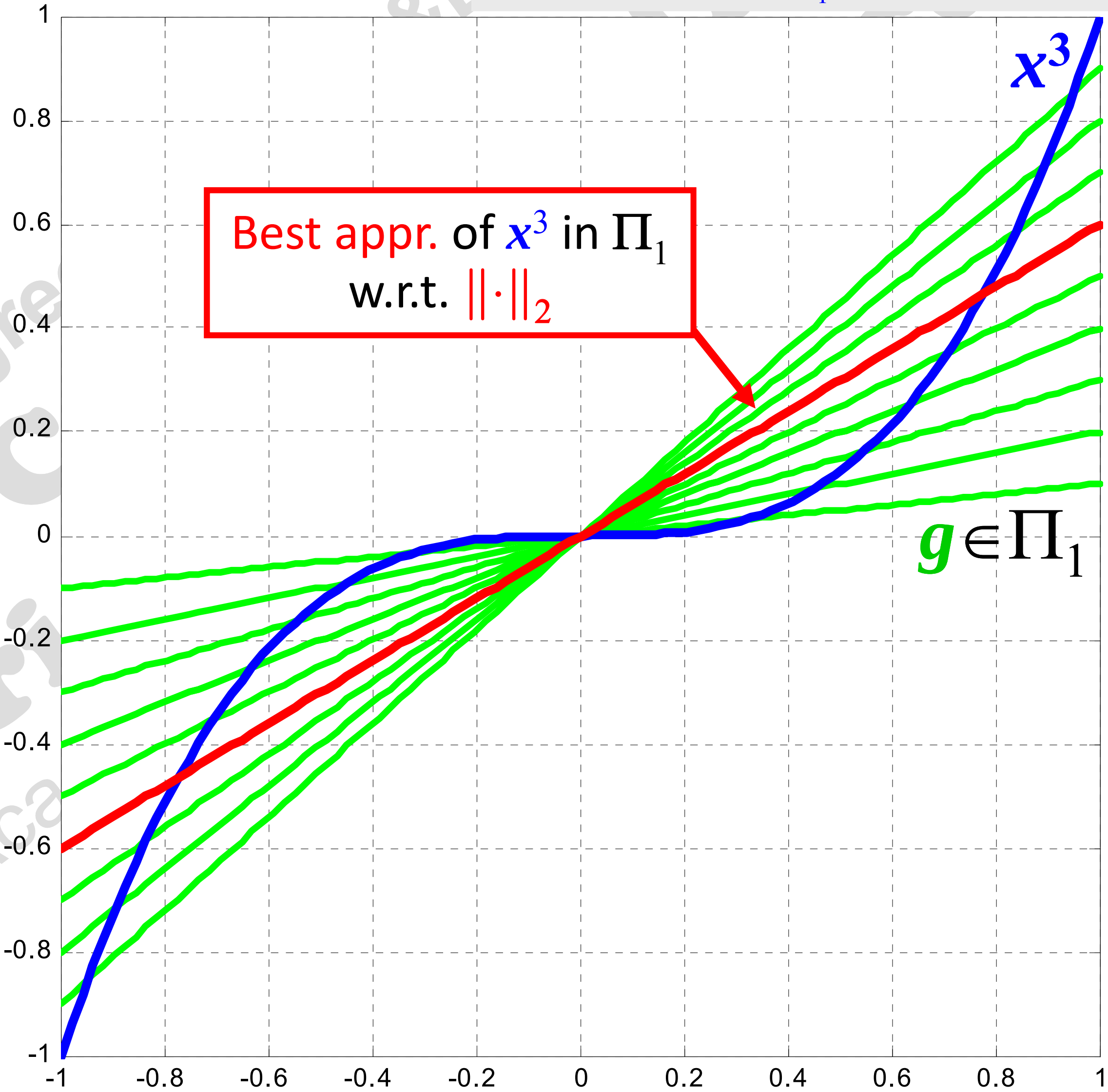
$$f^*(x) = c_1^* + c_2^* x \quad \Rightarrow \quad \left\|x^3 - f^*\right\|_2 = \min_{g\in\Pi_1}\left\|x^3 - g\right\|_2 \qquad (\Diamond)$$

$$(\Diamond) \qquad \left\|x^3 - g\right\|_2^2 = \int_{-1}^{+1}\left[x^3 - g(x)\right]^2 \mathrm{d}x$$

$$\Leftrightarrow \begin{cases} c_1^* = 0 \\ c_2^* = \dfrac{3}{5} \end{cases}$$

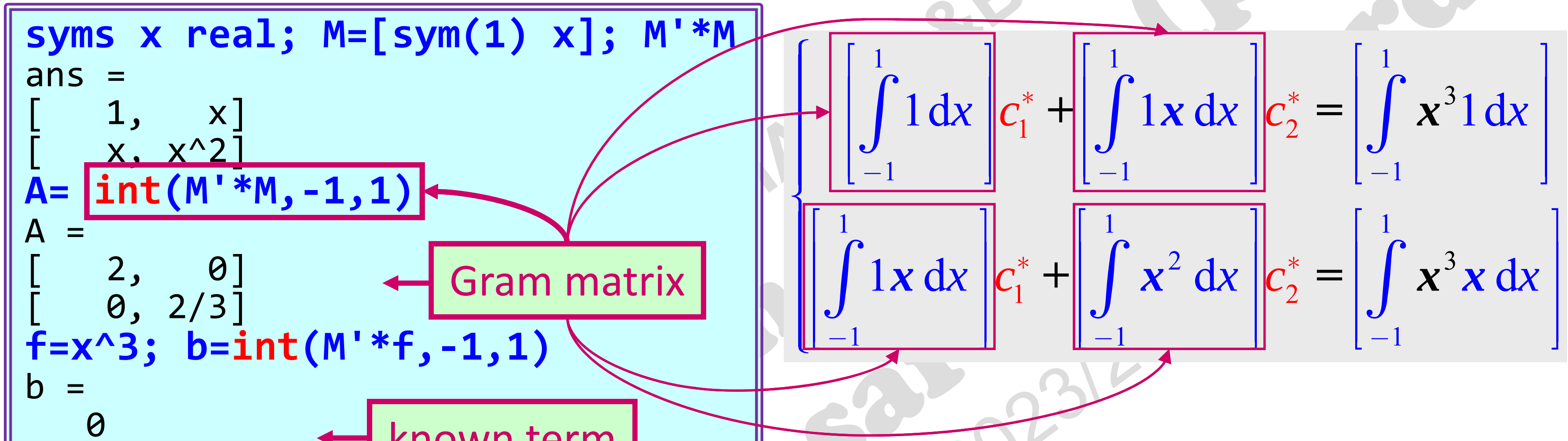
$$\Downarrow$$

$$\boxed{f^*(x) = \frac{3}{5}x}$$



Best appr. of x^3 in Π_1
w.r.t. $\|\cdot\|_2$

The System of Normal Equations is numerically created and solved; but it can also be described, more effectively for teaching goals, by means of the MATLAB Symbolic Math Toolbox:



```
syms x real; M=[sym(1) x]; M'*M
ans =
[   1,   x]
[   x, x^2]
A= int(M'*M,-1,1)
A =
[   2,   0]
[   0, 2/3]
f=x^3; b=int(M'*f,-1,1)
b =
   0
 2/5
c = A \ b
c =
  0
3/5
fstar = M*c
fstar  =
(3*x)/5
```

Gram matrix

known term

$$\begin{cases}\boxed{\left[\int_{-1}^{1} 1\,dx\right]} c_1^* + \boxed{\left[\int_{-1}^{1} 1x\,dx\right]} c_2^* = \left[\int_{-1}^{1} x^3 1\,dx\right]\\ \boxed{\left[\int_{-1}^{1} 1x\,dx\right]} c_1^* + \boxed{\left[\int_{-1}^{1} x^2\,dx\right]} c_2^* = \left[\int_{-1}^{1} x^3 x\,dx\right]\end{cases}$$

$$Ac^* = b^*$$

$$A = \begin{pmatrix} \int_{-1}^{1} 1\,dx & \int_{-1}^{1} 1x\,dx \\ \int_{-1}^{1} 1x\,dx & \int_{-1}^{1} x^2\,dx \end{pmatrix}, \quad b^* = \begin{pmatrix} \int_{-1}^{1} x^3 1\,dx \\ \int_{-1}^{1} x^3 x\,dx \end{pmatrix}$$

$$f^*(x) = c_1^* + c_2^* x = \frac{3}{5}x$$

$f^*(x)=\frac{3}{5}x$ is the **best approximation** of x^3 in [-1,+1] w.r.t. $\|\cdot\|_2$

Let us check the property of best approximation

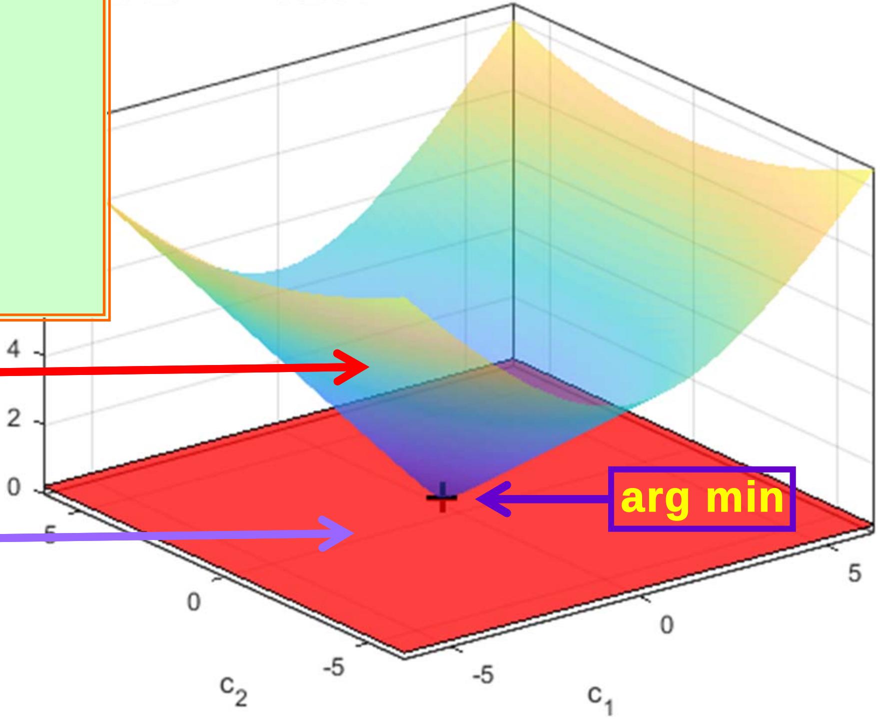
$$\left\|x^3-\frac{3}{5}x\right\|_2 \le \left\|x^3-(c_1+c_2x)\right\|_2 \;\forall c_1,c_2 \Longleftrightarrow \sqrt{\int_{-1}^{1}\left[x^3-\frac{3}{5}x\right]^2 \mathrm{d}x} \le \sqrt{\int_{-1}^{1}\left[x^3-(c_1+c_2x)\right]^2 \mathrm{d}x}$$

er_min

er

```
syms x c1 c2 real
er_min=sqrt(int((f-fstar)^2,-1,1));
er=sqrt(int((f-(c1+c2*x))^2,-1,1));
ezsurf(er); hold on; AX=axis;
surf(2*pi*[-1 1;-1 1],2*pi*[-1 -1;1 1], ...
            double(er_min)*ones(2,2))
axis(AX)
G=gradient(er); S=solve(G)
S = struct with fields:
    c1: 0
    c2: 3/5
```

arg min



`er=sqrt(int((x^3-(c1+c2*x))^2,-1,1))`

`er_min=sqrt(int((f-fstar)^2,-1,1))`

Example 3

Compute the best linear approximation $f^*(x)$ of $f(x)=x^3$ w.r.t. $\|\cdot\|_2$ in the subspace $\mathbf{P}_1[-1,+1]$ of 2nd degree trigonometric polynomials over [–1,+1]: $M_n = \mathbf{P}_1[-1,1] = \mathbf{span}\{1, \cos x, \sin x\}$

$$f^*(x) = c_1^* + c_2^* \cos x + c_3^* \sin x$$

unknowns

$$\begin{cases} \langle \varphi_1,\varphi_1\rangle c_1^* + \langle \varphi_1,\varphi_2\rangle c_2^* + \langle \varphi_1,\varphi_3\rangle c_3^* = \langle f,\varphi_1\rangle \\ \langle \varphi_2,\varphi_1\rangle c_1^* + \langle \varphi_2,\varphi_2\rangle c_2^* + \langle \varphi_2,\varphi_3\rangle c_3^* = \langle f,\varphi_2\rangle \\ \langle \varphi_3,\varphi_1\rangle c_1^* + \langle \varphi_3,\varphi_2\rangle c_2^* + \langle \varphi_3,\varphi_3\rangle c_3^* = \langle f,\varphi_3\rangle \end{cases}$$

Normal Equations

$$\begin{cases} \left[\int_{-1}^{1} 1\,dx\right] c_1^* + \left[\int_{-1}^{1} 1\cdot\cos x\,dx\right] c_2^* + \left[\int_{-1}^{1} 1\cdot\sin x\,dx\right] c_3^* = \left[\int_{-1}^{1} x^3 1\,dx\right] \\ \left[\int_{-1}^{1} \cos x\cdot 1\,dx\right] c_1^* + \left[\int_{-1}^{1} \cos^2 x\,dx\right] c_2^* + \left[\int_{-1}^{1} \cos x \sin x\,dx\right] c_3^* = \left[\int_{-1}^{1} x^3 \cos x\,dx\right] \\ \left[\int_{-1}^{1} \sin x\cdot 1\,dx\right] c_1^* + \left[\int_{-1}^{1} \sin x\cdot\cos x\,dx\right] c_2^* + \left[\int_{-1}^{1} \sin^2 x\,dx\right] c_3^* = \left[\int_{-1}^{1} x^3 \sin x\,dx\right] \end{cases}$$

$$f^*(x) = c_1^* + c_2^* \cos x + c_3^* \sin x$$

Normal Eqs solved by means of MATLAB Symbolic Math Toolbox:

$$Ac^* = b^*$$

$$A = \begin{pmatrix} \int_{-1}^{1} 1\,dx & \int_{-1}^{1} 1\cdot\cos x\;dx & \int_{-1}^{1} 1\cdot\sin x\;dx \\ \int_{-1}^{1} \cos x\cdot 1\;dx & \int_{-1}^{1} \cos^2 x\;dx & \int_{-1}^{1} \cos x \sin x\;dx \\ \int_{-1}^{1} \sin x\cdot 1\;dx & \int_{-1}^{1} \sin x\cdot\cos x\;dx & \int_{-1}^{1} \sin^2 x\;dx \end{pmatrix},\quad b^* = \begin{pmatrix} \int_{-1}^{1} x^3 1\;dx \\ \int_{-1}^{1} x^3 \cos x\;dx \\ \int_{-1}^{1} x^3 \sin x\;dx \end{pmatrix}$$

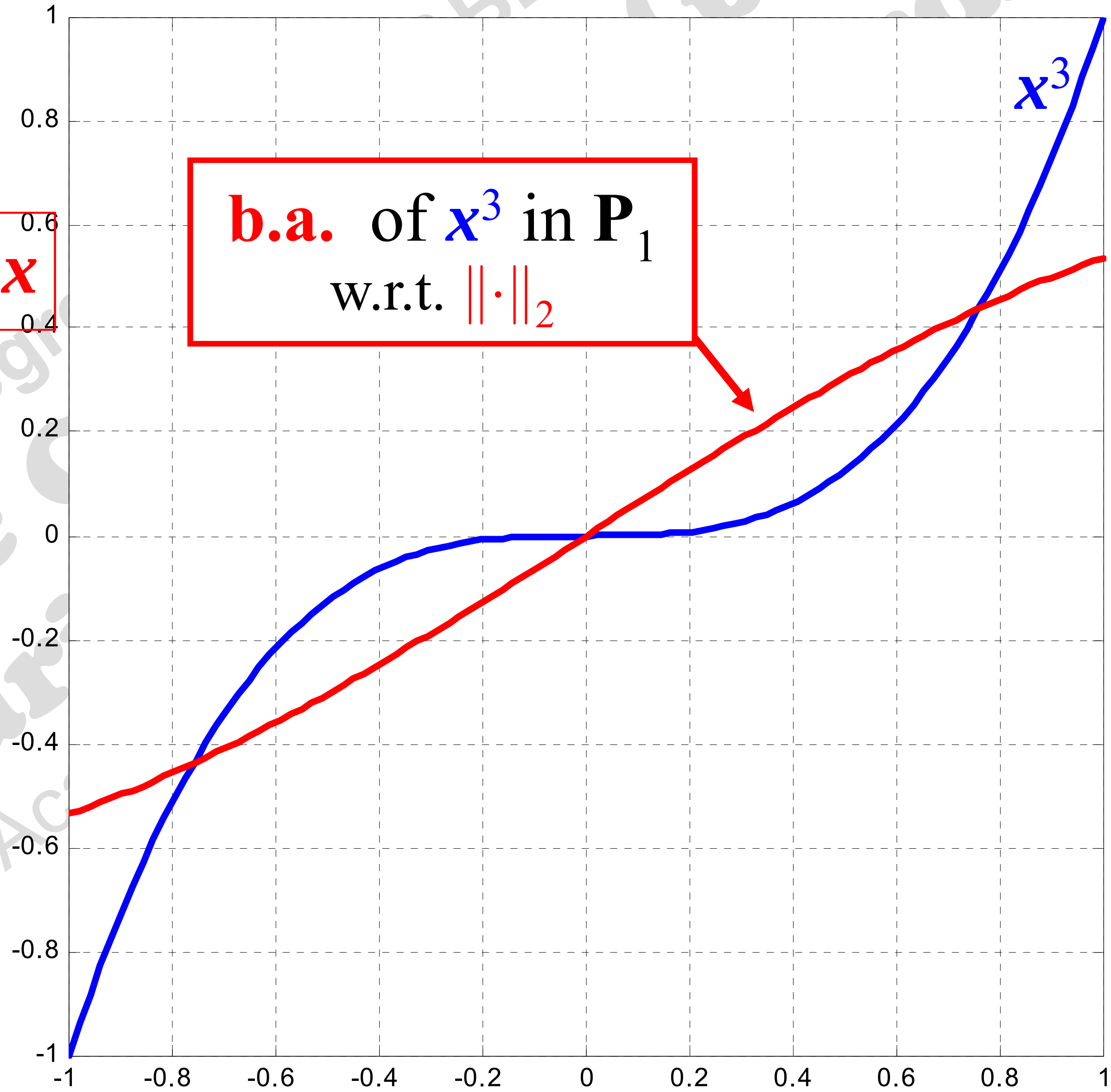
```
syms x real; M=[sym(1) cos(x) sin(x)];
A=int(M'*M,-1,1);  b=int(M'*x^3,-1,1);
c=A\b
c =
0
0
2*(-5*cos(1)+3*sin(1))/(cos(1)*sin(1)-1)
double(c)
ans =
         0
         0
    0.6495
```

$$f^*(x) = 0.6495\sin(x)$$

$$f^*(\mathbf{x}) = c_1^* + c_2^* \cos \mathbf{x} + c_3^* \sin \mathbf{x}$$

$$\left\| x^3 - f^* \right\|_2 = \min_{g \in \Pi_1} \left\| x^3 - g \right\|_2$$

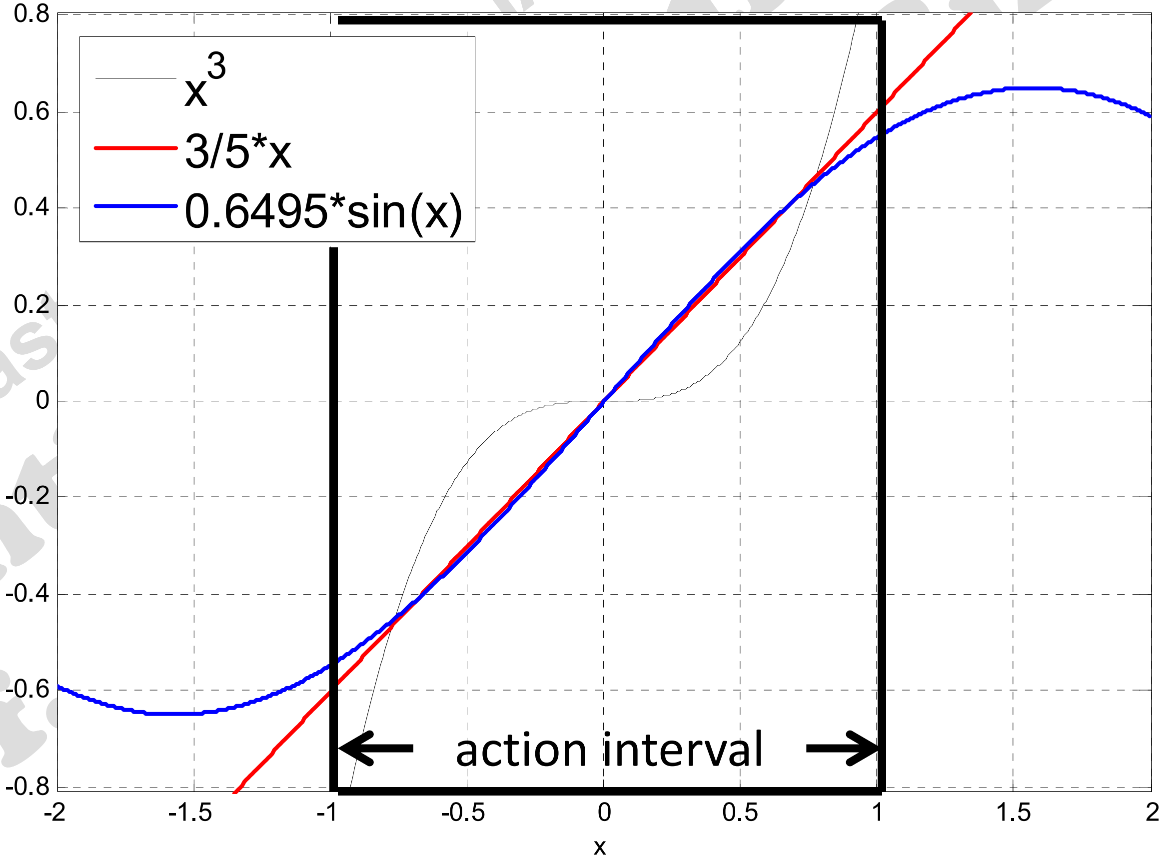
$$f^*(\mathbf{x}) \approx 0.6495 \sin \mathbf{x}$$



b.a. of x^3 in $\mathbf{P}_1$ w.r.t. $\|\cdot\|_2$

Exercise

Compare, from a “graphical” point of view, the last two approximations of the function $f(x)=x^3$ in $[-1,+1]$:



what can you say?

Contents

- **More on Normal Equations.**

To detect $f^*=\Sigma_k c^*\varphi_k$, the best approximation of f, w.r.t. $\|\cdot\|_2$, on the subspace

$$\boldsymbol{M}=\text{span}\{\varphi_1, \varphi_2 , \varphi_3, \ldots, \varphi_n\}$$

the Least Squares Method solves the Normal Equations

$\boldsymbol{x}^*$ LS sol.

$$\boldsymbol{G}\,\boldsymbol{x}^* = \boldsymbol{q}$$

$$G=\begin{pmatrix} \langle\varphi_1,\varphi_1\rangle & \langle\varphi_1,\varphi_2\rangle & \cdots & \langle\varphi_1,\varphi_n\rangle \\ \langle\varphi_2,\varphi_1\rangle & \langle\varphi_2,\varphi_2\rangle & \cdots & \langle\varphi_2,\varphi_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\varphi_n,\varphi_1\rangle & \langle\varphi_n,\varphi_2\rangle & \cdots & \langle\varphi_n,\varphi_n\rangle \end{pmatrix} \qquad q=\begin{pmatrix} \langle f,\varphi_1\rangle \\ \langle f,\varphi_2\rangle \\ \vdots \\ \langle f,\varphi_n\rangle \end{pmatrix}$$

where $\boldsymbol{G}$ is the Gram matrix of the subspace basis.

This generic form of the Normal Eqs applies to any normed Linear Space (containing vectors of $\mathbb{R}^n$, or functions), while the matrix form only holds for overdetermined linear systems.

The Normal Equations become simpler to be solved, if the basis vectors $\{\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)\}$ are:

orthogonal $\Leftrightarrow$ $\langle \varphi_k, \varphi_j \rangle \begin{cases} = 0 & k \neq j \\ \neq 0 & k = j \end{cases}$

Normal Eqs

$$\begin{pmatrix} \|\varphi_1\|_2^2 & & 0 \\ & \ddots & \\ 0 & & \|\varphi_n\|_2^2 \end{pmatrix} \underline{c}^* = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{pmatrix} \Rightarrow \boxed{c_j^* = \frac{\langle f, \varphi_j \rangle}{\|\varphi_j\|_2^2}}$$

diagonal matrix

solution of the Normal Equations

orthonormal $\Leftrightarrow$ $\langle \varphi_k, \varphi_j \rangle \begin{cases} = 0 & k \neq j \\ = 1 & k = j \end{cases}$

Normal Eqs

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \underline{c}^* = \begin{pmatrix} \langle f, \varphi_1 \rangle \\ \vdots \\ \langle f, \varphi_n \rangle \end{pmatrix} \Rightarrow \boxed{c_j^* = \langle f, \varphi_j \rangle}$$

identity matrix

named as **generalized Fourier coefficients of f in M_n**

We can use the Gram-Schmidt Orthonormalization algorithm.

Example: Find the linear b.a. f^* of $f(\boldsymbol{x})=\boldsymbol{x}^3$ w.r.t. $\|\cdot\|_2$ in the subspace $\Pi_1[-1,1]$ of 1st degree algebraic polynomials.

$$\Pi_1[-1,1]=\textbf{span}\boxed{\{1,\,x\}}$$

non-orthonormal basis

GSO algorithm

orthonormal basis $\left\{\frac{1}{\sqrt{2}},x\sqrt{\frac{3}{2}}\right\}$

$$c_j^*=\langle f,\varphi_j\rangle$$

$$f^*(\boldsymbol{x})=c_1^*\frac{1}{\sqrt{2}}+c_2^*\boldsymbol{x}\sqrt{\frac{3}{2}}$$

$$\begin{cases} c_1^*=\left\langle \boldsymbol{f},\frac{1}{\sqrt{2}}\right\rangle=\left\langle x^3,\frac{1}{\sqrt{2}}\right\rangle=\int_{-1}^{1}\frac{x^3}{\sqrt{2}}\,\mathrm{d}x=0 \\ c_2^*=\left\langle \boldsymbol{f},x\sqrt{\frac{3}{2}}\right\rangle=\left\langle x^3,x\sqrt{\frac{3}{2}}\right\rangle=\int_{-1}^{1}x^3x\sqrt{\frac{3}{2}}\,\mathrm{d}x=\frac{\sqrt{6}}{5} \end{cases}$$

$$f^*(\boldsymbol{x})=0\frac{1}{\sqrt{2}}+\frac{\sqrt{6}}{5}\boldsymbol{x}\sqrt{\frac{3}{2}}$$

$$\boxed{f^*(\boldsymbol{x})=\frac{3}{5}\boldsymbol{x}}$$

The solution is the same as in Example 2, obtained now without solving the Normal Eqs

Exercise

Find the linear best approximation f^* of $f(\boldsymbol{x})=\boldsymbol{x}^3$ w.r.t. $\|\cdot\|_2$ on the subspace $\mathbf{P}_1[-\pi,+\pi]$ of the 2nd degree trigonometric polynomials in $[-\pi,+\pi]$ by orthonormalizing, at first, the subspace basis $\{1,\cos(x),\sin(x)\}$, and compare the obtained solution to that one in Example 3 (SC2_11c.pdf): are they equal? Why not?

In the particular case of incompatible linear systems, solved by Least Squares method, the Gram matrix is

$$G = A^{\mathsf{T}} A$$

Consequences of orthonormalization

Gram-Schmidt Orthonormalization $\Longleftrightarrow$ $A=QR$ $\Rightarrow$

1. Orthogonal Projection Matrix P

$$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$$

for any columns in A

$\longrightarrow$

$$P = Q\,Q^{\mathsf{T}}$$

for orthonormal columns

2. Normal Eqs and QR factorization

3. Conditioning of Normal Equations

Consequences of orthonormalization

1. Orthogonal Projection Matrix

already seen in Linear Mappings

2. Normal Eqs and QR factorization

A=QR

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

$$(QR)^\top (QR)\, x^* = (QR)^\top b$$

$$R^\top \cancel{Q^\top Q} R\, x^* = R^\top Q^\top b$$

$$\cancel{R^\top} R\, x^* = \cancel{R^\top} Q^\top b$$

R is invertible if $rank(A)=n<m$

$$\boxed{R\, x^* = Q^\top b}$$

Upper triangular system
(simpler to solve)

Example

solve an incompatible system by means of the Least Squares Method and by orthonormalizing the basis

$$\begin{cases} x_1 + 2x_2 = 4 \\ x_1 + 5x_2 = 3 \\ 0x_1 + 0x_2 = 9 \end{cases} \Longleftrightarrow \boldsymbol{Ax=b} \qquad \boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 4 \\ 3 \\ 9 \end{pmatrix}$$

A=QR

$$\boldsymbol{Q} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 \end{pmatrix} \qquad \boldsymbol{R} = \begin{pmatrix} \sqrt{2} & 7/\sqrt{2} \\ & \frac{3}{2}\sqrt{2} \end{pmatrix}$$

Normal Equations $\boldsymbol{Rx} = \boldsymbol{Q^Tb}$

triangular system

$$\begin{pmatrix} \sqrt{2} & 7/\sqrt{2} \\ & \frac{3}{2}\sqrt{2} \end{pmatrix} x = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 9 \end{pmatrix}$$

The **solution is the same** as solving the Normal Eqs without orthonormalizing the basis (see Example 1)

$$\Rightarrow \boldsymbol{x} = \begin{pmatrix} \frac{14}{3} \\ -\frac{1}{3} \end{pmatrix}$$

Best Approximation in 2-norm

(prof. M. Rizzardi)

Consequences of orthonormalization

3. Conditioning of Normal Equations

In solving Normal Equations numerically, due to roundoff errors, it may happen that the solution is not that of Least Squares!

MATLAB example

```
A=[hilb(10); ones(1,10)*eps/1000];
b=ones(11,1);
[rank(A) rank([A b])]
ans =
    10    11
x1=A\b;
norm(b-A*x1)
ans =
    1
x2=(A'*A)\(A'*b);
norm(b-A*x2)
ans =
    1.00000008143431
```

Hilbert matrix is known to be ill-conditioned

incompatible system

Least Squares solution

minimum residual

Normal Equations solution

the solution of the Normal Eqs **does not reach the minimum** of $||\cdot||_2$ of the residual vector

If we orthonormalize the basis, by QR Factorization, now we get

```
A=[hilb(10); ones(1,10)*eps/1000];
b=ones(11,1);
[rank(A) rank([A b])]
ans =
    10    11
x1=A\b;
norm(b-A*x1)
ans =
    1
[Q,R]=qr(A,0);
x2=R\(Q'*b);
norm(b-A*x2)
ans =
    1
```

Hilbert matrix is known to be ill-conditioned

Least Squares solution

minimum residual

solution by QR factorization

the solution of the Normal Eqs **now reaches the minimum** of $\|\cdot\|_2$ of the residual vector

QR Factorization, in solving Normal Equations, did not amplify the data errors; thus we got the right solution!

In the case of an incompatible linear system ... remember that:

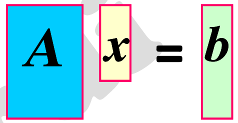
- The Least Squares solution can be also computed by means of **SVD Factorization**.
- The algorithm based on QR factorization is **more efficient** than that based on SVD factorization, but the latter is **numerically more stable**.

Normal Eqs and SVD factorization (*A*: full column rank)

The Least Squares solution x_{LS} of an overdetermined (incompatible) system $Ax=b$, as a solution of the Normal Equations, can be expressed by

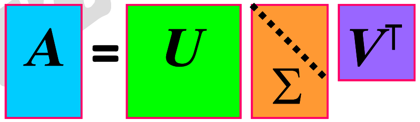
$$A^TA\,x_{LS} = A^Tb \quad \Rightarrow \quad x_{LS} = A^+b$$

$A_{m\times n}$, $m>n$



where A^+, of size $n\times m$, is the left pseudoinverse of A: $A^+ = (A^TA)^{-1}A^T$

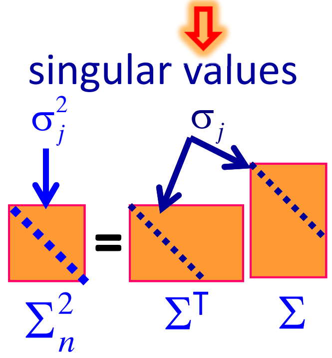
We want to express A^+ in terms of U, Σ, V, where $A = U\Sigma V^T$



$$x_{LS} = A^+b = \left(A^TA\right)^{-1}A^Tb = \left[\left(U\Sigma V^T\right)^T\left(U\Sigma V^T\right)\right]^{-1}\left(U\Sigma V^T\right)^T b =$$

$$= \left[V\Sigma^T\,\boxed{\cancel{U^TU}}\,\Sigma V^T\right]^{-1} V\Sigma^TU^Tb = \left[V\Sigma^T\Sigma V^T\right]^{-1} V\Sigma^TU^Tb =$$

$$= V\left(\Sigma^T\Sigma\right)^{-1}\boxed{\cancel{V^TV}}\,\Sigma^TU^Tb = V\left[\underset{\Sigma_n^2}{\boxed{\left(\Sigma^T\Sigma\right)}}^{-1}\Sigma^T\right]U^Tb = V\Sigma^+U^Tb$$



$$\Rightarrow \boxed{\boxed{\Sigma^+}} = \left(\Sigma^T\Sigma\right)^{-1}\Sigma^T = \left(\Sigma_n^2\right)^{-1}\Sigma^T = \underbrace{\begin{pmatrix} \sigma_1^{-2} & & & \\ & \sigma_2^{-2} & & \\ & & \ddots & \\ & & & \sigma_n^{-2} \end{pmatrix}}_{\left(\Sigma^T\Sigma\right)^{-1}} \overbrace{\left(\begin{array}{cccc|ccc} \sigma_1 & & & & 0 & \cdots & 0 \\ & \sigma_2 & & & 0 & \cdots & 0 \\ & & \ddots & & \vdots & \ddots & \vdots \\ & & & \sigma_n & 0 & \cdots & 0 \end{array}\right)}^{\Sigma^T} =$$

$$\Sigma_n$$

$$= \begin{pmatrix} \sigma_1^{-1} & & & & 0 & \cdots & 0 \\ & \sigma_2^{-1} & & & 0 & \cdots & 0 \\ & & \ddots & & \vdots & \ddots & \vdots \\ & & & \sigma_n^{-1} & 0 & \cdots & 0 \end{pmatrix} = \boxed{\left[\Sigma_n^{-1} \quad 0\right]}$$

$$\boxed{x_{LS} = \overbrace{\boxed{V\,\boxed{\Sigma^+}\,U^T}}^{A^+}b}$$

Example

solve an incompatible system by means of the Least Squares Method

Compare the LS solutions

rank(A)=2

$$A=\begin{pmatrix}1 & 2\\ 1 & 5\\ 0 & 0\end{pmatrix},\quad b=\begin{pmatrix}4\\ 3\\ 9\end{pmatrix}$$

1 $A^{\mathsf{T}}Ax=A^{\mathsf{T}}b$

solve the Normal Eqs

```
x1=(A'*A)\(A'*b)
x1 =
        4.6667
      -0.33333
```

unstable

2 $A=QR$

QR Factorization for Normal Equations

```
[Qn,Rn]=qr(A,0);
x2=Rn\(Qn'*b)
x2 =
        4.6667
      -0.33333
```

efficient

3 $A=USV^{\mathsf{T}}$

SVD Factorization for Normal Equations

```
[U,S,V]=svd(A,0);
x3=V*diag(1./diag(S))*U'*b
x3 =
        4.6667
      -0.33333
```

stable

the same solution

Compare LS solutions

```
n=11; A=vander(1:n); A=[A;rand(5,n)]; b=ones(size(A,1),1); r=rank(A);
disp([r rank([A b])])
     11    12
tic; xLS=A\b; tLS=toc; % LS solution
tic; xNE=(A'*A)\(A'*b); tNE=toc; % Normal Equations
tic; [Q,R]=qr(A);      xQR1=R\(Q'*b);    tQR1=toc; % QR 1
tic; [Qn,Rn]=qr(A,0); xQR2=Rn\(Qn'*b); tQR2=toc; % QR 2
tic; [U1,S1,V1]=svd(A); % SVD 1
xSVD1=V1*[diag(1./diag(S1(1:r,1:r))) zeros(r,m-r)]*U1'*b; tSVD1=toc;
tic; [U2,S2,V2]=svd(A,'econ'); % SVD2 equivalent to [U,S,V]=svd(A,0), since m>n
xSVD2=V2*diag(1./diag(S2))*U2'*b; tSVD2=toc;
format long
fprintf('\nnorm(b-A*xLS) = ');  disp(norm(b-A*xLS))
fprintf( 'norm(b-A*xQR1) = ');  disp(norm(b-A*xQR1))
fprintf( 'norm(b-A*xQR2) = ');  disp(norm(b-A*xQR2))
fprintf( 'norm(b-A*xSVD1) = '); disp(norm(b-A*xSVD1))
fprintf( 'norm(b-A*xSVD2) = '); disp(norm(b-A*xSVD2))
fprintf( 'norm(b-A*xNE) = ');   disp(norm(b-A*xNE))
format short g
fprintf('\ntime Least Square : %e',tLS)
fprintf('\ntime QR1 factoriz : %e',tQR1)
fprintf('\ntime QR2 factoriz : %e',tQR2)
fprintf('\ntime SVD1 factoriz: %e',tSVD1)
fprintf('\ntime SVD2 factoriz: %e',tSVD2)
fprintf('\ntime Normal Eqs.  : %e',tNE)
```

Vandermonde matrix is known to be ill-conditioned

$$A(16\times 11)$$

Use: MATLAB

tic

…

T=toc

to get the elapsed time

+ stable

– stable

+ efficient

– efficient

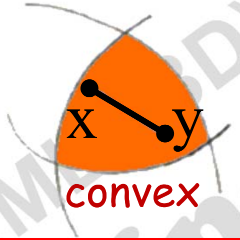
```
norm(b-A*xLS)   = 0.397654939767875
norm(b-A*xQR1)  = 0.397654939767937
norm(b-A*xQR2)  = 0.397654939767607
norm(b-A*xSVD1) = 0.397654939767366
norm(b-A*xSVD2) = 0.397654939767366
norm(b-A*xNE)   = 0.397654945772732

time Least Square : 1.608609e-04
time QR1 factoriz : 2.161868e-04
time QR2 factoriz : 1.717342e-04
time SVD1 factoriz: 9.626067e-04
time SVD2 factoriz: 4.384498e-04
time Normal Eqs.  : 7.707250e-04
```

Recall from p.1

Properties of X_{LS}, the set of LS solutions

$$X_{LS} = \{x^* \in \mathbb{R}^n : \|Ax^* - b\|_2 = \min \le \|Ay - b\|_2 \ \forall y \in \mathbb{R}^n\}$$

LS = Least Squares



A convex set contains all the segments between elements in the set.
A closed set contains all its limit points.

1. X_{LS} is a convex and closed set

2. $x^* \in X_{LS}$ ⇨ x^* is a solution of Normal Equations, i.e. its residual vector is orthogonal to $\mathscr{R}(A)$.

3. The system of Norm. Eqs $A^\top Ax = A^\top b$ admits one and only one solution, if rank($A^\top A$) is maximum (rank($A^\top A$)=n); otherwise, if rank(A)=$r<n$, the system is underdetermined.

4.

$$\exists! x_{LN} \in X_{LS} : \|x_{LN}\|_2 = \min\left\{\|x\|_2, \forall x \in X_{LS}\right\}$$

LN = Least Norm

There is only one solution of Normal Eqs of minimum $\|\cdot\|_2$ (x_{LN} is said the least norm solution), and it is the only element of X_{LS} belonging to $\mathscr{N}(A^\top A)^\perp = \mathscr{R}(A^\top A)$.

This is a particular case of the Problem of the Solution with Least Euclidean norm of an underdetermined system (see: SC2_06_NEW).

Solution with Minimum Euclidean norm of underdetermined Normal Eqs $A^TAx=A^Tb$, $A(m\times n)$, rank$(A)<\min\{m,n\}$

$x\in X_{LS} \quad x = \boxed{x_p} + x_n \ : \ \boxed{A^TAx_p=A^Tb} \wedge x_n\in\mathcal{N}(A^TA)$

particular solution

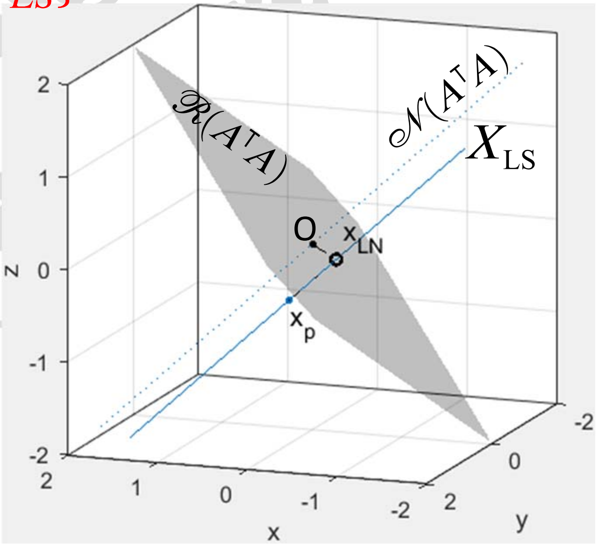
$x_p = x_r + x'_n,\ x_r\in\mathcal{R}(A^TA) : \|x_r\|_2 = \min\{\|x\|_2, x\in X_{LS}\}$

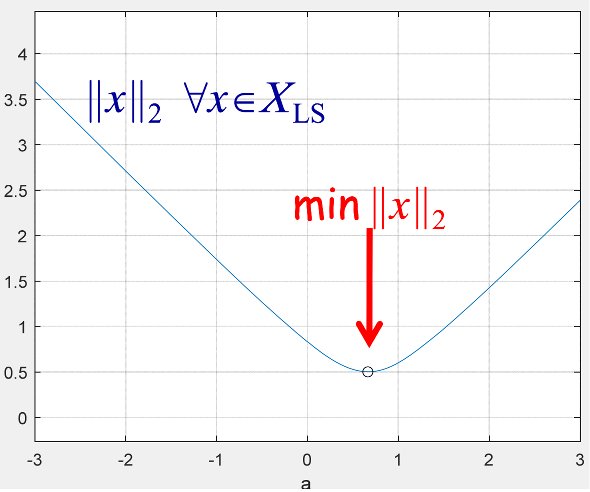
Example

```
A=[1 2 3 4; 5 6 7 8]';
A=[A  A(:,1)+A(:,2)];
b=[1 0 1 0]';
disp([rank(A) rank([A b])])
      2     3  % incompatible system A*x=b
disp([size(A'*A) rank(A'*A)])
        3      3      2  % non-max rank system A'*A*x = A'*b
xp=A\b;              % particular solution of Normal Eqs
M=A'*A; y=A'*b;  % underdetermined system (of Normal Eqs)
N=null(M);           % basis for the Null Space
syms a real; xn=N*a;  % Null Space
X=xp+xn;            % general solution of Normal Eqs
RMT=orth(M);        % orthonormal basis of ℛ(AᵀA)
P=RMT*RMT';         % orthogonal projection matrix
Pxp=P*xp;           % projection of xp onto ℛ(AᵀA)
xLN=pinv(M)*y;  % Moore-Penrose inverse
disp([norm(xLN) norm(Pxp) norm(xp)])
     0.50166   =   0.50166   <  0.83217
disp([norm(A*xLN-b) norm(A*Pxp-b) norm(A*xp-b)])
      0.89443   =   0.89443    =    0.89443
...
```

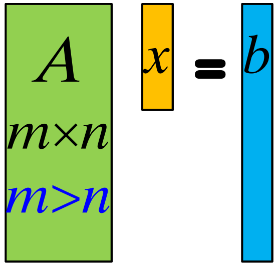
$\min\|x\|_2$

all LS solutions = $\arg\min\|Ax-b\|_2$





Solve incompatible systems in MATLAB



rank(A)=n ⇒ full (col) rank

```
A=[1 2 3 4; 5 6 7 8]';   b=[1 0 1 0]';
disp([rank(A) rank([A b])])
     2    3        % incompatible system A*x=b
disp(size(A))  % rank(A)=2 max
     4     2
```

∃ only one LS solution

$$\min \| Ax - b \|_2$$

the same solution

```
xBS=A\b  % LS solution with r=rank(A) non-zero components
xBS =
        -0.45
         0.25
```

```
xEN=(A'*A)\(A'*b) % Normal Equations solution
xEN =
        -0.45
         0.25
```

```
xLN=pinv(A)*b  % Moore-Penrose pseudoinverse
xLN =          % min||·||2  solution
        -0.45
         0.25
```

```
[U,S,V]=svd(A,'econ'); r=rank(S);
d=U'*b; xSVD=S(1:r,1:r)\d(1:r);
xSVD=V(:,1:r)*xSVD
xSVD =
        -0.45
         0.25
```

reduced SVD decomposition

```
[Q,R]=qr(A,0); r=rank(R); Qr=Q(:,1:r);
Rr=R(1:r,:); xQR1=Rr\(Qr'*b)
xQR1 =
        -0.45
         0.25
```

economy size QR decomposition

```
[Q,R,p]=qr(A,0); d=Q'*b; r=rank(R);
[m,n]=size(A);
xQR2=R(1:r,1:r)\d(1:r);
xQR2(r+1:n)=0;   xQR2(p)=xQR2
xQR2 =
        -0.45
         0.25
```

QR with a permutation vector p: A(:,p)=Q*R

Solve incompatible systems in MATLAB

$A_{m\times n}\,x = b$, $m>n$

rank(A)=$r<n$ ⇒ rank-deficient

```
A=[1 2 3 4;5 6 7 8]'; A=[A A(:,1)+A(:,2)]; b=[1 0 1 0]';
disp([rank(A) rank([A b])])
      2    3        % incompatible system A*x=b
disp(size(A)) % rank(A)=2 non-max
      4     3
```

infinitely many LS solutions

different LS solutions

the same solutions

```
xBS=A\b  % LS solution with rank(A) non-zero components
xBS =         0
            0.7
          -0.45
```

particular solution

```
xNE=(A'*A)\(A'*b) % Normal Equations solution
xNE  =  NaN
      ? Inf
       -Inf
```

```
[L,U,P]=lu(A'*A);  LU decomposition with partial pivoting
w=L\(P*A'*b); xp=U(1:r,:)\w(1:r)
xp =
              0
            0.7
          -0.45
```

```
xLN=pinv(A)*b  % min||·||2 solution
xLN =   -0.38333
         0.31667
       -0.066667
```

$\min\|Ax-b\|_2$

```
disp(norm(A*xBS-b))
     0.894427
disp(norm(A*xp-b))
     0.894427
disp(norm(A*xLN-b))
     0.894427
```

$\min\|x\|_2$

```
disp(norm(xBS))
     0.832166
disp(norm(xp))
     0.832166
disp(norm(xLN))
     0.501664
```

```
[U,S,V]=svd(A,'econ'); r=rank(S);
d=U'*b; xSVD=S(1:r,1:r)\d(1:r);
xSVD=V(:,1:r)*xSVD
xSVD =  -0.38333
         0.31667
       -0.066667
```

reduced SVD decomposition

```
[Q,R]=qr(A,0); r=rank(R); Qr=Q(:,1:r);
Rr=R(1:r,:); xQR1=Rr\(Qr'*b)
xQR1  =        0
             0.7
           -0.45
```

economy size QR decomposition

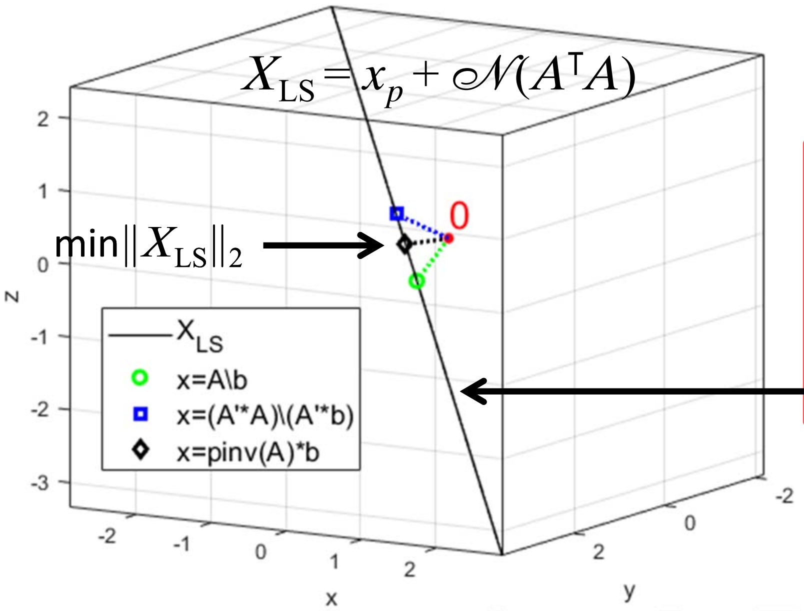
```
[Q,R,p]=qr(A,0);
d=Q'*b; r=rank(R); [m,n]=size(A);
xQR2=R(1:r,1:r)\d(1:r);
xQR2(r+1:n)=0;  xQR2(p)=xQR2
xQR2  =        0
             0.7
           -0.45
```

QR with permutation vector p: A(:,p)=Q*R

Solve incompatible systems in MATLAB

rank(A)=$r<n$ ⇒ rank-deficient

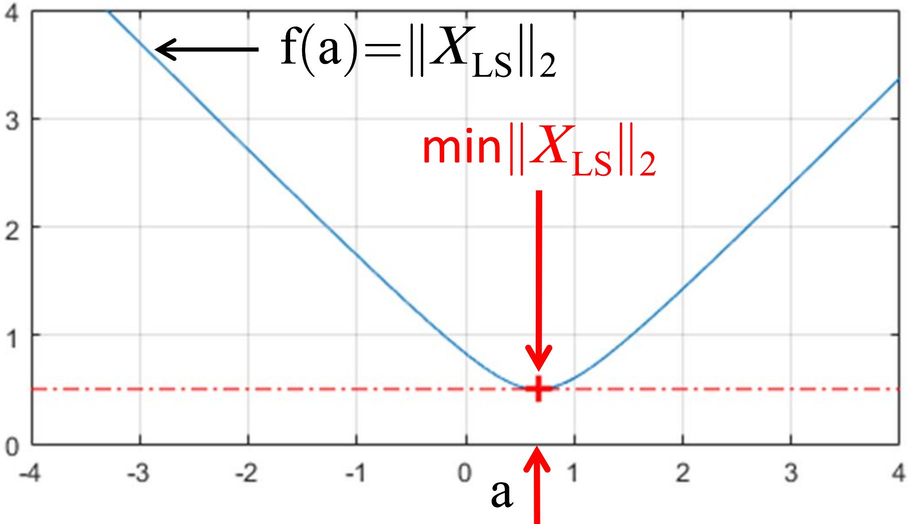
⇨ infinitely many LS solutions



```
nNE=null(A'*A); % 𝒩 (Gram matrix)
syms a real
N=nNE*a;
X = xp + N;
ezplot3(X(1),X(2),X(3),[-5 5])
```

X_{LS}

```
fplot(simplify(norm(X)),[-4 4])
```



amin

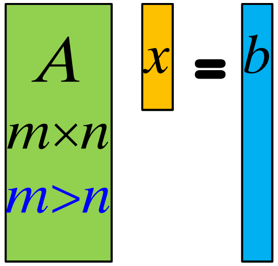
```
xLN=pinv(A)*b; % Moore-Penrose inverse
amin=solve(diff(norm(X),a)); % argmin
amin=double(amin)
amin =
      0.66395
Ymin=double(subs(norm(X),a,amin));
disp([Ymin norm(xLN)])
       0.50166    =    0.50166
```

$\min\|X_{LS}\|_2$

norm of the solution
from Moore-Penrose
pseudoinverse

Solve incompatible systems in MATLAB

rank(A)=$r<n \Rightarrow$ rank-deficient



```
u=[1 1 1 1]'; v=[1 -1 1 -1]'; A=u*v'; A=A(:,1:3); b=[1 0 1 0]';
disp([rank(A) rank([A b])])
      1    2         % incompatible system A*x=b
disp(size(A))  % rank(A)=2 non-max
      4      3
```

infinitely many LS solutions

the same solutions

```
xBS=A\b  % LS solution with rank(A) non-zero components
xBS =      0.5
             0
             0
```

```
disp(norm(A*xBS-b))
     1
```

```
r=rank(A); [L,U,P]=lu([A'*A A'*b]);
xLU=U(1:r,1:n)\U(1:r,n+1)
xLU =
       0.5
         0
         0
```

LU decomposition
with partial pivoting

```
xEN=(A'*A)\(A'*b) % Normal Equations solution
xEN  = NaN
   ?   NaN
       NaN
```

```
xLN=pinv(A)*b  % min‖·‖2 solution
xLN  =  0.16667
       -0.16667
        0.16667
```

```
disp(norm(A*xLN-b))
     1
```

```
[U,S,V]=svd(A,'econ'); r=rank(S);
d=U'*b; xSVD=S(1:r,1:r)\d(1:r);
xSVD=V(:,1:r)*xSVD
xSVD =  0.16667
       -0.16667
        0.16667
```

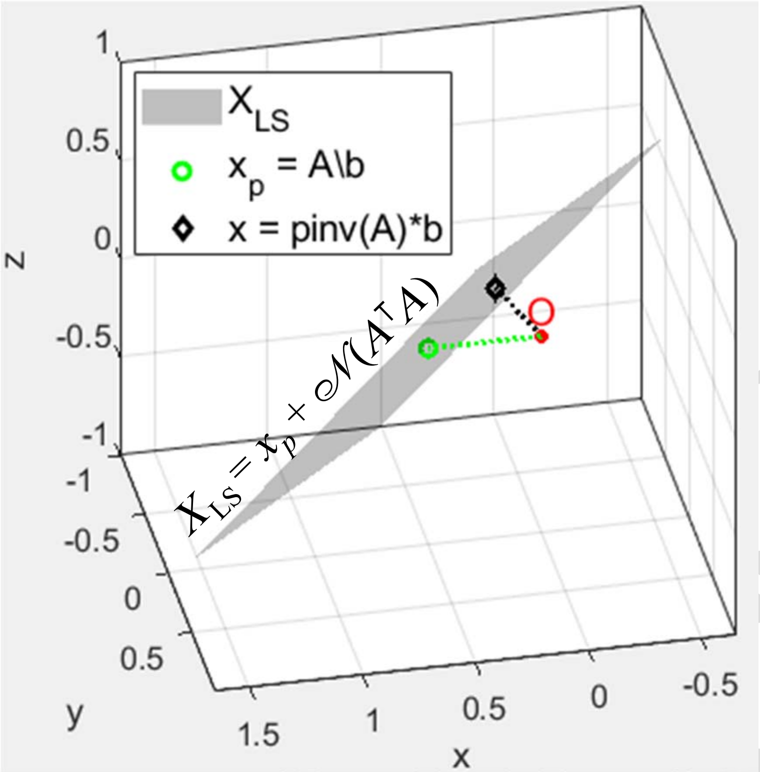
reduced SVD
decomposition

```
[Q,R]=qr(A,0); r=rank(R); Qr=Q(:,1:r);
Rr=R(1:r,:); xQR1=Rr\(Qr'*b)
xQR1 =     0.5
             0
             0
```

economy size QR
decomposition

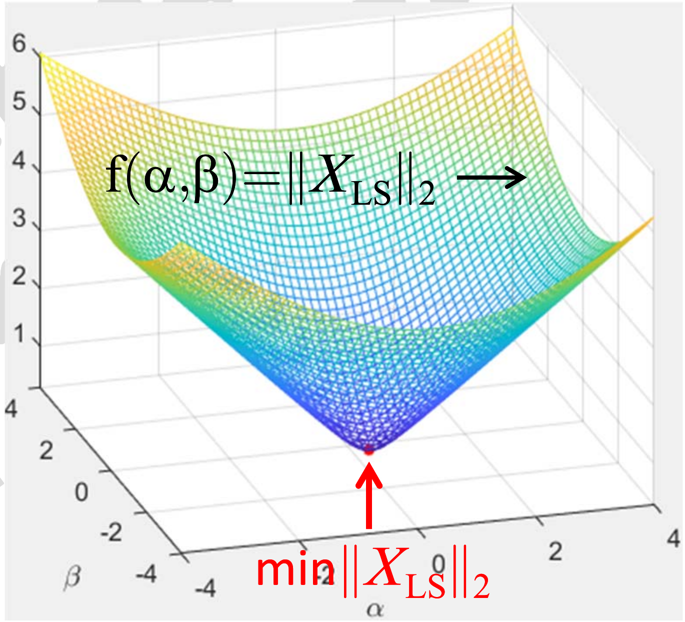
```
[Q,R,p]=qr(A,0); d=Q'*b; r=rank(R);
[m,n]=size(A);
xQR2=R(1:r,1:r)\d(1:r);
xQR2(r+1:n)=0;  xQR2(p)=xQR2
xQR2 =     0.5
             0
             0
```

QR with permutation
vector p: A(:,p)=Q*R



```
nEN=null(A'*A); % 𝒩 of Gram matrix
syms alfa beta real
N=nEN*[alfa;beta];
X=xLU + N;
fsurf(X(1),X(2),X(3),[-1 1])
```

X_{LS}



Exercise

Verify, by means of MATLAB Symbolic Math Toolbox, that $\min\|X_{LS}\|_2$ is reached by the LS solution computed by **`pinv()`**.

Exercise

Find the best fit circle Γ of a sample of N data in $\mathbb{R}^2$. What is the Least Norm ($\min\|\cdot\|_2$) solution? Is the solution unique? If so, why?

The unknown circle Γ, centered at (a,b) and of radius R, has equation:

$$(x - a)^2 + (y - b)^2 = R^2$$

similarly for the eq.: $x^2 + y^2 + \alpha x + \beta y + \gamma = 0$

We want that all the samples (x_i,y_i) belong to the circle:

$$\forall i=1,\dots,N \qquad (x_i - a)^2 + (y_i - b)^2 = R^2$$

i.e.

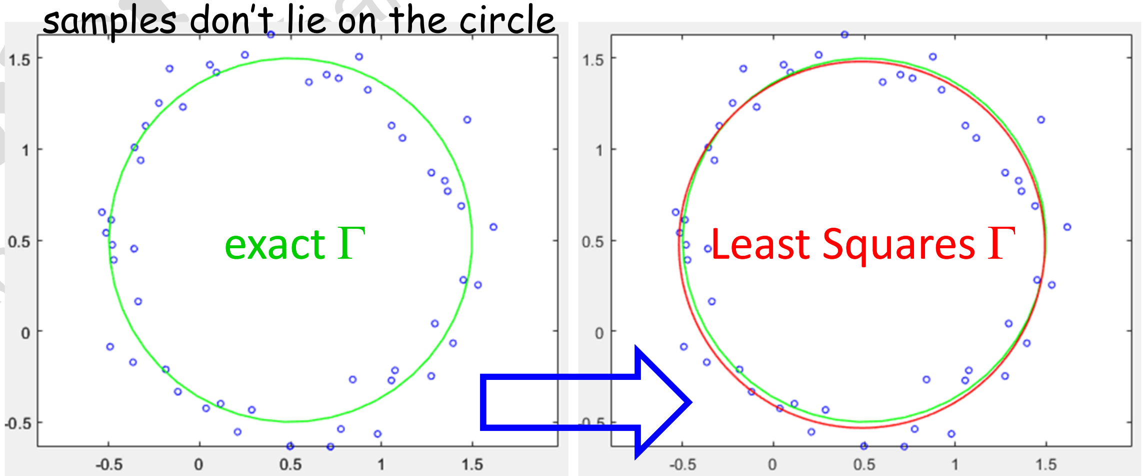
$$x_i^2 + y_i^2 + a^2 + b^2 - 2x_ia - 2y_ib = R^2$$

By reordering unknowns, we get: $\forall i=1,\dots,N \qquad 2x_ia + 2y_ib + R^2 - a^2 - b^2 = x_i^2 + y_i^2$

i.e. $[2x_i \;\; 2y_i \;\; 1]\begin{pmatrix} a \\ b \\ R^2 - a^2 - b^2 \end{pmatrix} = x_i^2 + y_i^2$ incompatible system

```
X0=0.5; Y0=0.5; r0=1;
N=50; %=,100,150,200 num. of samples
t=linspace(-pi,pi,N)';
Xi=X0+r0*cos(t); % exact Γ
Yi=Y0+r0*sin(t);
perc=0.15;    percentage of noise
xi=Xi+perc*(2*rand(N,1)-1);
yi=Yi+perc*(2*rand(N,1)-1);
plot(xi,yi,'ob'); axis equal
hold on; plot(Xi,Yi,'r')
```

samples don't lie on the circle



Solve also the problem by using the eq. $x^2 + y^2 + \alpha x + \beta y + \gamma = 0$

Exercise: wind tunnel experiment

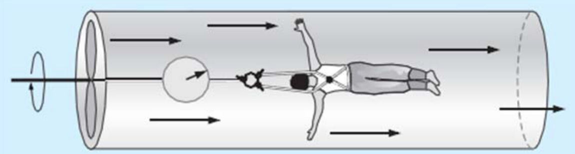
Solve by means of Linear Least Squares the following fitting problem*:

Numerical problem: linear LS fitting
Statistical problem: linear regression

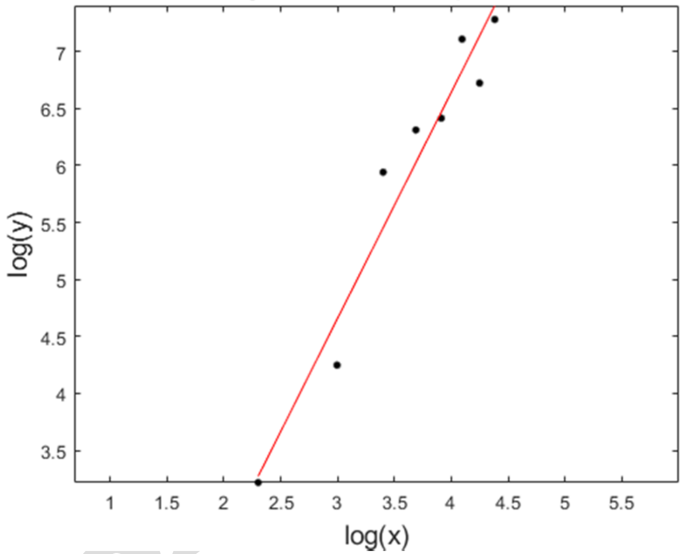
speed v (m/s): `v=[10 20 30 40 50 60 70 80]';`
force F (N): `F=[25 70 380 550 610 1220 830 1450]';`

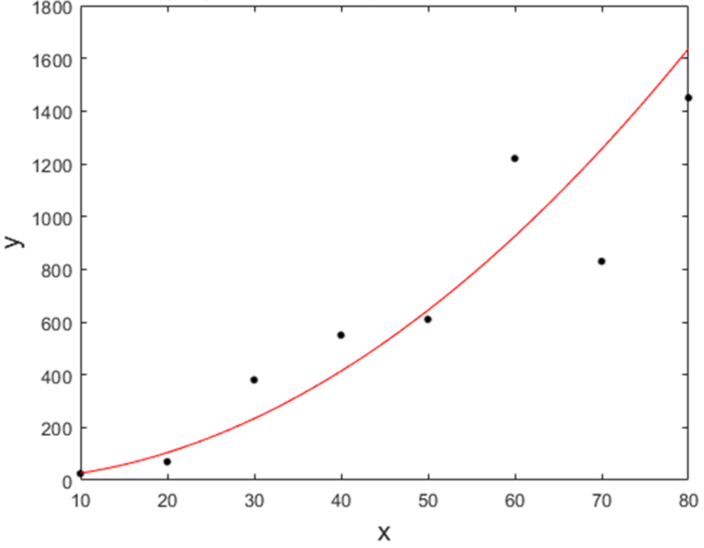
modeled by power function: $y = f(x) = \boxed{a\,x^b,}\quad a,b\in\mathbb{R}$

* Wind tunnel experiment: how the air resistance force depends on wind speed



The fitting model is non-linear, but it can be simply linearized by applying the logarithmic transformation.





Attention! The “log trick” could lead to a solution that differs from the desired. For example: if the residual y_i-$f(x_i)$ has a normal distribution, $\log(y_i)$-$\log[f(x_i)]$ has not.

Contents

- **Best linear approximation w.r.t. $\|\cdot\|_2$: the case of infinite dimension subspaces.**
- **Concept of convergence in norm.**

Best linear approximation w.r.t. $\|\cdot\|_2$ in a subspace

- **finite dimension subspace**

Theor. → discrete case / continuous case

Theorem for existence and uniqueness

- **infinite dimension subspace**

no Theor. $\cancel{\rightarrow}$ discrete case / continuous case ?

Let us consider the **residual error** of the best linear approximation in a finite dimensional subspace: how does the residual error perform as the subspace dimension increases?

Let

$f_n^*(x)$: the best approx. of $\boldsymbol{f}(x)$ w.r.t. $\|\cdot\|_2$ on $\boldsymbol{M}_n=\mathrm{span}\{\varphi_1,\varphi_2,\dots,\varphi_n\}$ where φ_k are orthonormal; dim $\boldsymbol{M}_n=n$

$f_{n+1}^*(x)$: the best ap. of $\boldsymbol{f}(x)$ w.r.t. $\|\cdot\|_2$ on $\boldsymbol{M}_{n+1}=\mathrm{span}\{\varphi_1,\varphi_2,\dots,\varphi_n,\varphi_{n+1}\}$ $(\boldsymbol{M}_{n+1}\supset \boldsymbol{M}_n)$ where φ_k are orthonormal. dim $\boldsymbol{M}_{n+1}=n+1$

$$\| f(x) - f_{n+1}^*(x) \|_2 \quad \textbf{?} \quad \| f(x) - f_n^*(x) \|_2$$

Given $f_n^*(x)$, for computing $f_{n+1}^*(x)$, we don't need to repeat all the calculations, but we only need to compute the last coefficient $\boldsymbol{c}_{n+1}^*=\langle f,\varphi_{n+1}\rangle$ in the linear combination:

$$f^*_{n+1} = \sum_{k=1}^{n+1} c_k^*\varphi_k = \underbrace{\sum_{k=1}^{n} c_k^*\varphi_k}_{f_n^*} + c_{n+1}^*\varphi_{n+1}$$

The two residual errors are such that:

$$\boxed{\left\|f-f_{n+1}^*\right\|_2^2 \le \left\|f-f_n^*\right\|_2^2}$$

Proof:

Since $f_n^* = \sum_{k=1}^{n} c_k^* \varphi_k$ ⇒

$$\boxed{\left\|f_n^*\right\|_2^2} = \left\langle f_n^*, f_n^* \right\rangle = \sum_{k=1}^{n}\sum_{h=1}^{n} c_k^* c_h^* \overbrace{\left\langle \varphi_k, \varphi_h \right\rangle}^{\delta_{ij}} = \sum_{k=1}^{n}\left|c_k^*\right|^2$$

⇒

$$\boxed{\left\langle f, f_n^* \right\rangle} = \sum_{k=1}^{n} c_k^* \underbrace{\left\langle f, \varphi_k \right\rangle}_{c_k^*} = \sum_{k=1}^{n}\left|c_k^*\right|^2$$

orthonormal system $\{\varphi_k\}$

we have $\forall n$

$$\boxed{\left\|f-f_n^*\right\|_2^2} = \left\langle f-f_n^*, f-f_n^* \right\rangle = \left\|f\right\|_2^2 + \boxed{\left\|f_n^*\right\|_2^2} - 2\boxed{\left\langle f, f_n^* \right\rangle} =$$

$$= \left\|f\right\|_2^2 + \sum_{k=1}^{n}\left|c_k^*\right|^2 - 2\sum_{k=1}^{n}\left|c_k^*\right|^2 =$$

$$\boxed{= \left\|f\right\|_2^2 - \sum_{k=1}^{n}\left|c_k^*\right|^2}$$

$$\left\|f-f_{n+1}^*\right\|_2^2 = \left\|f\right\|_2^2 - \sum_{k=1}^{n+1}\left|c_k^*\right|^2 = \boxed{\left\|f\right\|_2^2 - \sum_{k=1}^{n}\left|c_k^*\right|^2} - \left|c_{n+1}^*\right|^2 = \boxed{\left\|f-f_n^*\right\|_2^2} - \left|c_{n+1}^*\right|^2$$

$$f_n^* = \sum_{k=1}^{n} c_k^* \varphi_k$$ the best approximation of f in M_n $\forall n$

$$\Rightarrow \boxed{\left\| f - f_{n+1}^* \right\|_2^2 \le \left\| f - f_n^* \right\|_2^2} \not\Rightarrow \lim_{n\to\infty} \underbrace{\left\| f - f_n^* \right\|_2}_{\text{residual}} = 0$$

If we have an infinite orthonormal system of basis functions

$$\{\varphi_k(x)\}_{k=1,\,\ldots,\,\infty}$$

then the sequence of residual errors $\{\|\boldsymbol{f}(x)-\boldsymbol{f}_n^*(x)\|_2\}_n$ in the best approximations $\{\boldsymbol{f}_n^*(x)\}_n$ is non-increasing ($\le$), …but this doesn't imply that it is decreasing ($<$) and infinitesimal (res.$\to 0$).

QUESTIONS

- What happens if $\dim \boldsymbol{M}_n = n \longrightarrow \infty$?
- Is it possible that the sequence of best approximations of $f(x)$ w.r.t. $\|\cdot\|_2$ converges in $\|\cdot\|_2$ to $f(x)$?

$$\lim_{n\to\infty} \left\| f - f_n^* \right\|_2 = 0 \qquad \textbf{?}$$

$\forall n\ f_n^*(x) =$ best approx. of $f(x)$ in $\boldsymbol{M}_n$=span$\{\varphi_1,\varphi_2,\ldots,\varphi_n\}$ w.r.t. $\|\cdot\|_2$

$\{\varphi_k\}_k$ orthonormal basis

$$\overset{?}{\Longrightarrow} \qquad \boxed{\lim_{n\to\infty}\left\|f - f_n^*\right\|_2 = 0}$$

IT DOESN'T HAPPEN AUTOMATICALLY

We have to add more assumptions to the Linear Space $\boldsymbol{X}$ containing $f(x)$, and to the orthonormal basis of $\boldsymbol{X}$ ($\{\varphi_k\}_{k=1,2,\ldots,\infty}$):

- $\boldsymbol{X}$ must be a **Hilbert Space** (*complete metric space*).
- $\{\varphi_k\}_{k=1,2,\ldots,\infty}$ must be a **complete** orthonormal system w.r.t. $\|\cdot\|_2$ in $\boldsymbol{X}$.

$\forall n\ f_n^*(x) =$ best approx. of $f(x)$ in $\boldsymbol{M}_n$=span$\{\varphi_1,\varphi_2,\ldots,\varphi_n\}$ w.r.t. $\|\cdot\|_2$

$\{\varphi_k\}_k$ orthonormal basis

From the previous proof, we get:

$$0 \le \|f - f_n^*\|_2^2 = \|f\|_2^2 - \sum_{k=1}^{n} |c_k^*|^2 \qquad \forall n$$

This always implies that:

$$\boxed{\sum_{k=1}^{n} |c_k^*|^2 \le \|f\|_2^2 \qquad \forall n} \Longrightarrow \boxed{\sum_{k=1}^{\infty} |c_k^*|^2 \le \|f\|_2^2}$$

Bessel's inequality

$\le\ \forall\{f_n^*\}$, not $=$

If, in addition, $\{\varphi_k\}_k$ is complete in the *Hilbert Space* $\boldsymbol{X}$, then the following holds:

$$\sum_{k=1}^{\infty} |c_k^*|^2 = \|f\|_2^2$$

Parseval's equality $\Longleftrightarrow$ $\lim_{n\to\infty} \|f - f_n^*\|_2 = 0$

convergence in $\|\cdot\|_2$

Parseval's Theorem represents the generalization of Pythagoras' Theorem for right triangles in spaces with ∞ dimensions.

It can be proved that the ***trigonometric functions***

$$\boxed{\left\{\frac{1}{\sqrt{2\pi}},\left\{\frac{\cos(kx)}{\sqrt{\pi}},\frac{\sin(kx)}{\sqrt{\pi}}\right\}_k\right\}}$$

or, equivalently, the ***exponential functions***

$$\boxed{\left\{\frac{e^{ikx}}{\sqrt{2\pi}}\right\}_k}$$

Euler's formula
$e^{i\theta} = \cos\theta + i\sin\theta$

form a ***complete orthonormal system*** w.r.t. $\|\cdot\|_2$ in the **Hilbert space** $L^2([-\pi,+\pi])$ of square integrable (or summable) functions over $[-\pi,+\pi]$.

This implies that the **Fourier Series** of $f(x) \in \boldsymbol{L}^2([-\pi,+\pi])$ **converges in mean square** or **in quadratic mean** (i.e., w.r.t. $\|\cdot\|_2$) to $f(x)$.

Convergence in norm

A sequence of functions $\{\varphi_n(x)\}$ is said to be **convergent in norm** to the function $\varphi(x)$, over an interval $[a,b]$, if $\forall x \in [a,b]$

$$\lim_n \|\varphi(x) - \varphi_n(x)\| = 0 \qquad \{\varphi_n(x)\} \to \varphi(x)$$

... by specifying the norm, we get

$$\lim_n \|\varphi(x) - \varphi_n(x)\|_\infty = 0 \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad \text{uniform convergence}$$

Example: if φ is an *analytic holomorphic function*, then the sequence of partial sums of its Taylor series (power series) is uniformly convergent to φ

$$\lim_n \|\varphi(x) - \varphi_n(x)\|_2 = 0 \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad \text{convergence in mean square (or quadratic mean convergence)}$$

Example: if φ is a *square integrable function*, then the sequence of partial sums of its Fourier series (trigonometric series) converges in quadratic mean to φ

$\{f_n{}^*\}_n$ convergence in norm $\overset{\text{def}}{\Longleftrightarrow}$ $\lim\limits_{n}\left\|f-f_n^*\right\|=0$

convergence in $\|\cdot\|_\infty$ $\Longrightarrow$ convergence in $\|\cdot\|_2$

(uniform convergence) $\not\Longleftarrow$ (convergence in mean square)

EXAMPLE: convergence in $\|\cdot\|_2$

In $\underset{[a,b]}{C^0([-1,+1])}$ the sequence of functions $\{f_n(x)\}_n$ $\quad f_n(x)=\sqrt{\dfrac{n}{1+n^4x^2}}$

converges in $\|\cdot\|_2$ to the zero function.

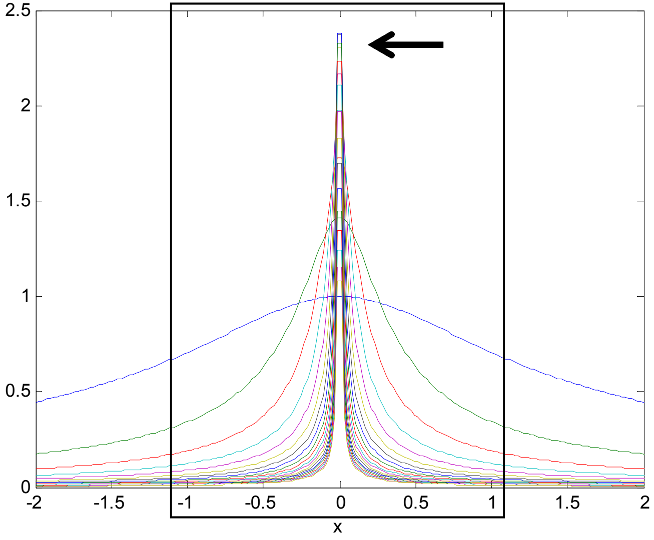
$$\|f\|_2^2=\int_a^b |f(x)|^2\,dx<\infty$$

In facts, it is $\|0-f_n(x)\|_2^2=\displaystyle\int_{-1}^{+1}\frac{n}{1+n^4x^2}\,\mathrm{d}x=\frac{2}{n}\arctan n^2\to 0\quad \text{per}\quad n\to\infty$

But it does not converge in $\|\cdot\|_\infty$ because it goes to ∞ at 0 (diverges)

$$f_n(0)=\sqrt{n}\to\infty\quad\text{per}\quad n\to\infty$$

$$d(f,g)=\sup_{x\in[-1,+1]}|f(x)-g(x)|=\|f-g\|_\infty$$



$\{f_n^*\}_n$ convergence in norm $\overset{\text{def}}{\Longleftrightarrow}$ $\lim\limits_n \left\|f - f_n^*\right\| = 0$

convergence in $\|\cdot\|_\infty$ $\underset{\not\Leftarrow}{\Rightarrow}$ convergence in $\|\cdot\|_2$

(uniform convergence) (convergence in mean square)

EXAMPLE: convergence in $\|\cdot\|_\infty$

In $C^0(\mathbb{R})$ the sequence of functions $\{f_n(x)\}_n$ $\quad f_n(x) = \dfrac{nx^4}{1+nx^2}$

converges uniformly (in $\|\cdot\|_\infty$) to the function x^2: in facts, it is

$$\lim_n \left\|x^2 - f_n(x)\right\|_\infty^2 = \lim_n \max_{x\in\mathbb{R}} \left|\underbrace{x^2 - \frac{nx^4}{1+n^4x^2}}\right| = 0$$

$$\lim_n \left|\frac{1+\cancel{nx^4} - \cancel{nx^4}}{1+nx^2}\right| = 0$$

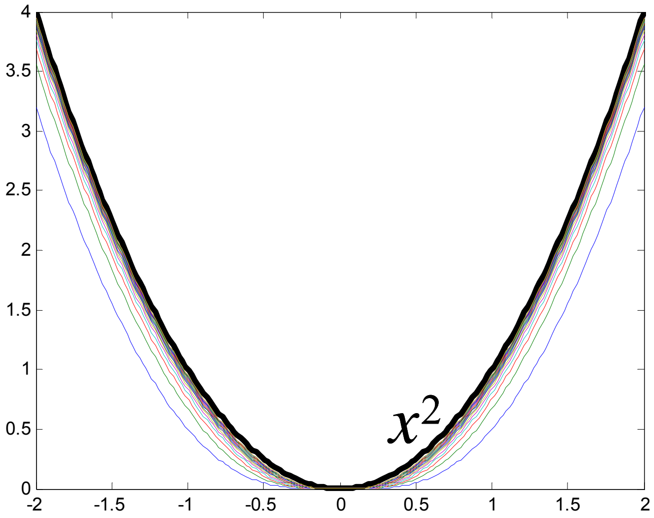
also converges in $\|\cdot\|_2$ because

for uniformly convergent sequences, the following holds:

$$\lim_n \| f_n(x) - x^2 \|_2 = \boxed{\lim_n} \int_{\mathbb{R}} | f_n(x) - x^2 | \, dx =$$

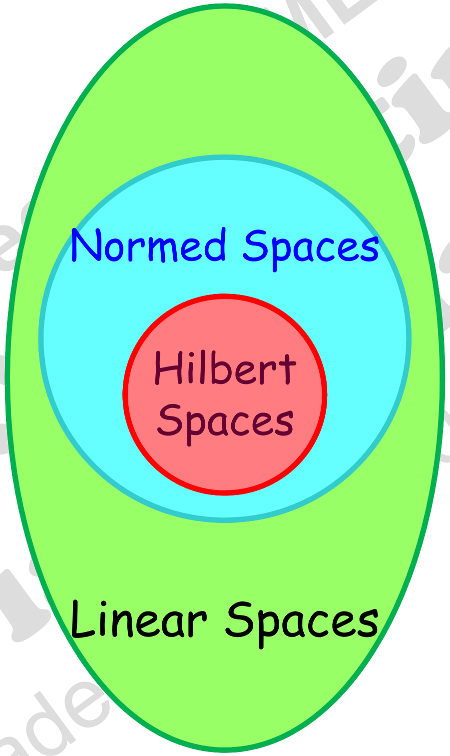
swap the 2 operators

$$= \int_{\mathbb{R}} \boxed{\lim_n} f_n(x) \, dx = \int_{\mathbb{R}} 0 \, dx = 0$$



The theorem for the existence and uniqueness of the best approx. w.r.t. $\|\cdot\|_2$ is only valid for *finite dimension* subspaces.

Hilbert Spaces are introduced to guarantee the existence of the best approx. w.r.t. $\|\cdot\|_2$ in any subspace, even of infinite dimension.



In practice, **Hilbert Spaces** make it possible to mantain, even in infinite-dimensional spaces, the same "geometry" as **Euclidean Linear Spaces** (i.e.: finite vector lenght, angle between vectors, Pythagoras' Theorem, …), which is familiar for spaces such as $\mathbb{R}^2$, $\mathbb{R}^3$, …, $\mathbb{R}^n$.