III Least-square data fitting

- A standard problem in numerical analysis is that of data fitting using least-square methods.

① Simple example using Choleski decomposition

Suppose we have a set of n points in $\mathbb{R}^2$

$$\{(x_i, y_i)\}_{i=1,n}$$

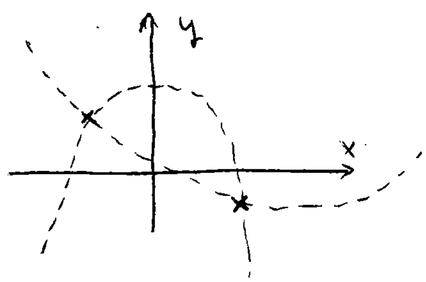
and we want to find the best fit to a parabola

$$y = a_2x^2 + a_1x + a_0 = f(x; a_0, a_1, a_2)$$

$\Rightarrow$ we need to find the closest solution to the set of equations

$$\begin{cases} y_1 = a_2x_1^2 + a_1x_1 + a_0 \\ \quad\vdots \\ y_n = a_2x_n^2 + a_1x_n + a_0 \end{cases}$$

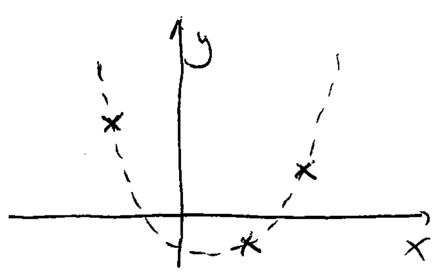
- If $n<3$ the problem is underspecified and the solution is not unique



Many parabola fit exactly through 2 pts.
$\rightarrow$ fit through parabola not informative

- if $n=3$ the problem usually has a unique solution, which is the solution of the linear system

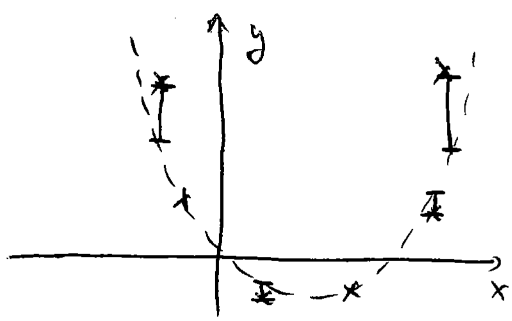
$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$



• If $n > 3$ the problem is overspecified:

$$\begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & & \\ \vdots & & \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

$\Rightarrow$ we do not seek an exact solution but the one which minimizes the least-square error (the norm of the error).



$\Rightarrow$ we want to minimize

$$E = \sum_{i=1}^{n} \left[y_i - (a_2 x_i^2 + a_1 x_i + a_0) \right]^2$$

$$= \sum_{i=1}^{n} \left[y_i - f(x_i; a_0, a_1, a_2) \right]^2$$

by choosing the best possible values of $\{a_0, a_1, a_2\}$.

If we rewrite the linear system above as

$$\underline{b} = M \underline{a} \quad \text{with} \quad \underline{b} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \underline{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & & \\ 1 & x_n & x_n^2 \end{pmatrix}$$

then the error $E = \| \underline{b} - M\underline{a} \|^2$

$$= (b - M\underline{a})^T (b - M\underline{a}).$$

$$= b^T b - a^T M^T b - b^T M a + a^T M^T M a$$

Minimizing the error implies finding solutions to

$$\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = \frac{\partial E}{\partial a_2} = 0$$

or equivalently $\nabla E = 0$ where $\nabla \equiv \nabla_a$.

$\Rightarrow$ $\nabla E = 0$ implies

$$M^T \underline{b} + \underline{b}^T M = M^T M \underline{a} + \underline{a}^T M^T M$$

Now note that if $M^T M \underline{a} = M^T \underline{b}$ then automatically

$\underline{b}^T M = \underline{a}^T M^T M$ so it suffices to solve

$$\boxed{M^T M \underline{a} = M^T \underline{b}}$$

Since M is an $(n \times 3)$ matrix, $M^T M$ is a (3×3) matrix; $\underline{a}$ and $\underline{b}$ are both 3-vectors so <u>this</u> system is well posed and should have a unique solution (unless M is singular).

$\Rightarrow$ Method of solution:

- Construct the system $M\underline{a} = \underline{b}$
- construct the system $\tilde{M}\underline{a} = \tilde{b}$ with $\tilde{M} = M^T M \qquad \tilde{b} = M^T b$
- solve for $\underline{a}$.

$\Rightarrow$ Since $\tilde{M} = M^T M$, $\tilde{M}^T = (M^T M)^T = M^T M = \tilde{M}$

$\Rightarrow$ $\tilde{M}$ is a symmetric matrix

Moreover, $\forall v$,

$$v^T \tilde{M} v = v^T M^T M v = (v^T M^T)(Mv)$$
$$= \| Mv \|^2 > 0$$

So the Cholesky method is an obvious candidate for solving this LS problem.

(2) <u>See sample program, data + fitted curve.</u>

For a cubic data fitting problem.

HW: How would you rewrite the program to make it versatile?

IV QR factorization

① Review on orthogonal matrices & vectors

Definitions

- For a real matrix A, we define the transpose of A
  $$A^T = \{ a_{ji} \} \quad \text{if} \quad A = \{ a_{ij} \}.$$
  (switching rows & columns).
- A is symmetric if $A = A^T$
- For a complex matrix A, we define the conjugate of A
  $$A^* = \{ a_{ji}^* \} \quad \text{if} \quad A = \{ a_{ij} \}$$
  ($*$ denotes complex conjugation).
- A is hermitian if $A = A^*$.

In what follows we consider real matrices only. See Trefethen & Bau for generalization to complex systems.

- The inner product of 2 vectors u and v is
  $$u^T v = \sum_{k=1}^{N} u_k v_k$$
- The Euclidean norm of a vector is $\|u\| = \sqrt{u^T u}$
- Two vectors are orthogonal if their inner product is null
- Two sets of vectors X and Y are orthogonal if every vector in X is perpendicular to every vector in Y.
- A set of vectors X is orthogonal if every vector in X is orthogonal to all the other ones

- A set of vectors X is <u>orthonormal</u>

$\forall u \in X \quad \forall v \in X$ with $v \neq u$

$$\underbrace{u^T v = 0}_{\text{mutual orthogonality}} \quad \text{and} \quad \underbrace{\|u\| = \|v\| = 1}_{\text{normalization}}.$$

- A matrix A is said to be orthogonal (unitary in the complex case) if

$$A^T A = A A^T = I \iff A^T = A^{-1}$$

$\iff$ the column vectors of A form an orthonormal set

② <u>Householder Matrices</u>

<u>Definition</u> Given a vector w, with $\|w\| = 1$ then

$P = I - 2ww^T$ is a Householder matrix

<u>Properties</u>

ⓐ P is a symmetric, orthogonal matrix

<u>Proof</u> • $P_{ij} = \delta_{ij} - 2w_i w_j$

$P_{ji} = \delta_{ji} - 2w_j w_i = P_{ij}$

$\rightarrow$ symmetric. ✓

• $PP^T = (I - 2ww^T)(I - 2ww^T)^T$

$= P^2 = I - 4ww^T + 4w\underbrace{w^T w}_{1}w^T$

$= I$

$\Rightarrow P^T = P^{-1}$ orthogonal ✓

(b) P is a reflection across the plane $\perp$ to w

Proof:

These are properties of a refection

- $Pw = (I - 2ww^T)w = w - 2w\underbrace{w^Tw}_{I}$
  $= -w$

  ↳ P changes w into $-w$.

- Now consider $u \perp$ to w then
  $Pu = (I - 2ww^T)u = u$
  since $w^Tu = 0$

  ↳ P project a vector $\perp$ to u onto itself

(3) QR factorization

- QR factorization consists in writing an input matrix A as

  $A = QR$ where Q is orthogonal
  R is upper triangular.

- Note that A need <u>not</u> be a square matrix:

  if A is $n \times m$
  Q is $n \times n$
  R is $n \times m$ $\Rightarrow$ $\begin{pmatrix} n\times m \end{pmatrix} = \begin{pmatrix} n \times n \end{pmatrix} \begin{pmatrix} n\times m \end{pmatrix}$

  (upper triangular still means $R_{ij} = 0$ if $i > j$

- Relationship between Householder matrices & QR factorization.

  (a) Let $\underline{a}_1$ be the first column vector of A

  Let $w_1^T = \dfrac{1}{\sqrt{2s_1(s_1 - a_{11})}} (a_{11} - s_1, a_{21}, \dots, a_{n1})$

  where $s_1 = \pm\|\underline{a}_1\|$ (sign chosen so that $a_{11} - s_1 \neq 0$)

then

$$w_1^T w_1 = \frac{1}{2s_1(s_1-a_{11})}\left[(a_{11}-s_1)^2 + a_{21}^2 + \dots + a_{n1}^2\right]$$

$$= \frac{1}{2s_1(s_1-a_{11})}\left[\underbrace{\underbrace{a_{11}^2+\dots+a_{n1}^2}_{\|\underline{a}_1\|^2} - 2s_1a_{11} + s_1^2}_{2s_1^2-2s_1a_{11} = 2s_1(s_1-a_{11})}\right]$$

$$= 1$$

let $P^{(1)} = I - 2w_1w_1^T$ be the Householder matrix constructed from w_1

Then

$$P^{(1)}\underline{a}_1 = \underline{a}_1 - 2\underline{w}_1\underline{w}_1^T\underline{a}_1$$

$$= \underline{a}_1 - 2w_1\left(\frac{(a_{11}-s_1)a_{11} + a_{21}^2 + \dots a_{n1}^2}{\sqrt{2s_1(s_1-a_{11})}}\right)$$

$$= \underline{a}_1 - 2\underline{w}_1\left(\frac{s_1^2 - s_1a_{11}}{\sqrt{2s_1(s_1-a_{11})}}\right)$$

$$= \underline{a}_1 - \sqrt{2s_1(s_1-a_{11})}\,\underline{w}_1$$

$$= \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} - \begin{pmatrix} a_{11}-s_1 \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} = \begin{pmatrix} s_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\rightarrow P^{(1)}$ acts on $\underline{a}_1$ to zero all coefficients except the first, which becomes a signed norm of $\underline{a}_1$

$$\Rightarrow P^{(1)}A = \begin{pmatrix} s_1 & x & x & \cdots & x \\ 0 & x & & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x & x & \cdots & x \end{pmatrix} = \tilde{A}$$

↑ a new matrix.

(b) Let's now define $w_2^T = \frac{1}{\sqrt{2s_2(s_2-\tilde{a}_{22})}}\left(0, \tilde{a}_{22}-s_2, \tilde{a}_{32} \dots \tilde{a}_{n2}\right)$

with $s_2 = \left(\sum_{k=2}^{n} \tilde{a}_{k2}^2\right)^{1/2}$

Then $w_2^T w_2 = 1$, so

Then $P^{(2)} = I - 2w_2 w_2^T$ is a Householder matrix

and

$$P^{(2)}\tilde{A} = \begin{pmatrix} s_1 & x & x & \cdots & x \\ 0 & s_2 & x & & \vdots \\ \vdots & 0 & x & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & x & \cdots & x \end{pmatrix}$$

↳ the action of $P^{(2)}$ is to leave the first line unchanged, to set $\tilde{a}_{22}$ to s_2 and zero out all coefficients in the second column below $\tilde{a}_{22}$.

(c) → If we do this repetitively then

$$P^{(n-1)} \cdots P^{(2)} P^{(1)} A = R$$ ← the successive operations turn A into an upper triangular matrix.

But since $P^{(i)}$ is an orthogonal matrix,

$\prod_i P^{(i)}$ is also orthogonal

→ let $Q^T = Q^{-1} = \prod_i P^{(i)}$ then

$A = QR$. as desired.

(4) Use of QR for least-square methods

In standard least-square solution, the over-determine

$n \times m$ problem

$$AX = B \qquad (n > m)$$

is reduced to

$$A^T A X = A^T B \quad \rightarrow \text{ an } (m \times m) \text{ problem.}$$

A common problem arises if the entries of A span several orders of magnitude ⇒ truncation errors accumulate in the calculation of $A^T A$