Least Squares Application

Given data pairs (x_i, y_i) where $i = 1, \ldots n$ we seek to minimize the error

$$E(\alpha, \beta) \equiv \sum_{i=1}^{n} \left(y_i - \alpha (x_i - \bar{x}) - \beta \right)^2$$

where

$$\bar{x} \equiv \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \text{(average)}$$

After some calculations solving the coupled eqns

$$(1) \qquad E_\alpha(\alpha, \beta) = 0$$

$$(2) \qquad E_\beta(\alpha, \beta) = 0$$

results in

$$(3) \qquad \alpha = \frac{\sum y_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \qquad \beta = \frac{1}{n} \sum_{i=1}^{n} y_i$$

Rather than solve the calculus problem (1)-(2) we pose it as a linear algebra problem.

For most data the linear fit problem

$$\underbrace{\begin{bmatrix} x_1 - \bar{x} & 1 \\ x_2 - \bar{x} & 1 \\ \vdots & \vdots \\ x_n - \bar{x} & 1 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{b}$$

$$\boxed{Ax = b}$$

is illposed (no solution)

Instead we seek to solve

$$\min_{x} \|Ax - b\|^2$$

over $x = (\alpha, \beta)^T$. The solution of this can be shown to be that given in (3)

Least Squares Solutions

Defn: Let $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$. We say x_0 is a least squares soln of $Ax = b$ if

$$F(x_0) = \min_x \|Ax - b\|^2$$

where $F(x) \equiv \|Ax - b\|^2$

Remarks

(1) The definition presumes the existence of a minimizer x_0 which we shall shortly prove.

(2) If $A \in \mathbb{R}^{n \times n}$ is invertible the $x_0 = A^{-1}b$ is the unique minimizer with $F(x_0) = 0$.

(3) Least squares solutions need not be unique, even if A is square. For example if $N(A) \neq 0$, choose $y \in N(A)$

$$\begin{aligned} F(x_0 + y) &= \|A(x_0 + y) - b\|^2 \\ &= \|Ax_0 + Ay - b\|^2 \\ &= \|Ax_0 - b\|^2 \\ &= F(x_0) \end{aligned}$$

so $x_0 + y$ least squares soln $\forall y \in N(A) \neq 0$.

EXISTENCE OF LEAST SQUARES SOLN(S)

To show a least squares solution we first let

$$b = b_r + b_\perp \qquad b_r \in R(A) \quad b_\perp = R(A)^\perp$$

Then

$$\|Ax - b\|^2 = \langle \underbrace{(Ax - b_r)}_{\text{in } R(A)} - b_\perp , (Ax - b_r) - b_\perp \rangle$$

Expandind and $(Ax - b_r) \perp b_\perp$ yields

$$(1) \quad \|Ax - b\|^2 = \|Ax - b_r\|^2 + \|b_\perp\|^2$$

Since $b_r \in R(A)$, $\exists x_0$ such that

$$Ax_0 = b_r$$

For any such x_0

$$\|Ax - b\|^2 = \|b_\perp\|^2$$

and th minimum of (1) in x is attained for $x = x_0$.

Normal Equations (Calculus Derivation)

Let x_0 minimize $F(x) = \|Ax - b\|^2$. Then

$$G(\varepsilon) \equiv F(x_0 + \varepsilon y) \qquad \varepsilon \in \mathbb{R}$$

is minimized at $\varepsilon = 0$ for any fixed $y \in \mathbb{R}^m$. Expanding $\|A(x_0 + \varepsilon y) - b\|^2$ yields

$$G(\varepsilon) = \|Ax_0 - b\|^2 + 2\varepsilon \langle y, A^*(Ax_0 - b)\rangle + \varepsilon^2 \|Ay\|^2$$

Necessarily $G'(0) = 0$ or

$$\langle y, A^*(Ax_0 - b)\rangle = 0 \qquad \forall y \in \mathbb{R}^m$$

Since y arbitrary x_0 must be a solution of

$$\boxed{A^*A\, x_0 = A^* b}$$

NORMAL EQUATIONS

Remarks

(1) Solutions exist

(2) Solutions may not be unique

Normal Equations (Existence)

By the Fredholm Alternative

$$A^*A x_0 = A^*b$$

has a solution $\Leftrightarrow$ $\langle A^*b, v\rangle = 0 \quad \forall v \in N(A^*A)$

$$A^*A v = 0$$

implies

(1) $\qquad A v \in R(A) \qquad\qquad A v \in N(A^*)$

But FAT implies $R(A) \perp N(A^*)$
so (1) implies

$$\langle Av, Av\rangle = 0$$

$$\|Av\|^2 = 0$$

$$v \in N(A)$$

Now we check the FAT solvability cond:

$$\begin{aligned}\langle A^*b, v\rangle &= \langle b, Av\rangle \\ &= \langle b, 0\rangle \\ &= 0\end{aligned}$$

Conclude

$$A^*A x = A^*b$$

has a soln.

Normal Equations/Moore Penrose Pseudoinverse

(1) $$A^*A\,x = A^*b$$

can have nonunique solutions:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad A^*A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

so that $N(A^*A) \neq 0$. But if (1) has a unique solution

$$x = A' b \quad \equiv (A^*A)^{-1}A^*$$

where A' is called the Moore Penrose Pseudoinverse.

Note if $A \in \mathbb{R}^{n\times n}$ and invertible

$$A' = A^{-1}(A^*)^{-1}A^* = A^{-1}I = A^{-1}$$

Theorem If $N(A) = 0$ then

$$A' = (A^*A)^{-1}A^*$$

exists and Norm. Eqns have unique soln

Pf/ Fredholm Alternative

$$(A^*A)\,x = A^*b \qquad \text{unique soln}$$

$$\Updownarrow$$

$$(A^*A)\,x = 0 \qquad \forall x$$

$$\Updownarrow \qquad\qquad \langle A^*Ax, x\rangle = 0 \quad \forall x$$

$$\|Ax\|^2 = 0 \quad \forall x$$

$$Ax = 0 \qquad \forall x$$

EXAMPLE Find the least squares fit (linear) through $(1,1)$, $(2,2)$, $(1,2)$

$$y_i = \alpha(x_i - \bar{x}) + \beta \qquad \bar{x} = \tfrac{4}{3}$$

$$A = \begin{bmatrix} -\frac{1}{3} & 1 \\ \frac{2}{3} & 1 \\ -\frac{1}{3} & 1 \end{bmatrix} \qquad A^* = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 1 & 1 & 1 \end{bmatrix} \qquad A^*A = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & 3 \end{bmatrix}$$

Since A^*A diagonal, easily inverted

$$(A^*A) = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}.$$

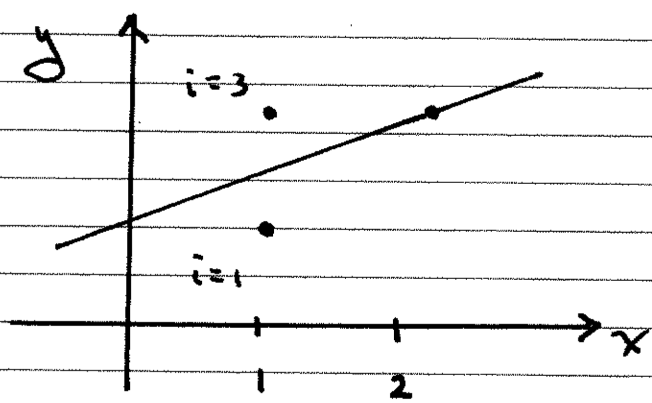
Compute pseudo inverse

$$A' = (A^*A)^{-1}A^* = \begin{bmatrix} -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

For our $Ax = b$ problem $x = (\alpha, \beta)^T$ and b is the column of y_i values, ... $b = (1,2,2)^T$

$$x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = A'b = \begin{pmatrix} \frac{1}{2} \\ \frac{5}{3} \end{pmatrix}$$

yields



$$y = \tfrac{1}{2}\left(x - \tfrac{4}{3}\right) + \tfrac{5}{3}$$

Normal Equations (Geometrical Derivation)

Want to minimize $F(x) \equiv \|Ax - b\|^2$.

Let

$$b = b_r + b_\perp \qquad b_r \in R(A) \quad b_\perp \in R(A)^\perp$$

Then, as before,

$$\|Ax - b\|^2 = \|Ax - b_r\|^2 + \|b_\perp\|^2$$

wlog must minimize first term $\|Ax - b_r\|^2$

① $$b_\perp \in R(A)^\perp \iff \langle b_\perp, Ax \rangle = 0 \quad \forall x$$

$$\iff \langle A^* b_\perp, x \rangle = 0 \quad \forall x$$

$$\Rightarrow b_\perp \in N(A^*)$$

② Since $b_r \in R(A)$, $\exists x$ s.t. $Ax = b_r$

$$Ax = b_r$$

$$Ax = b - b_\perp$$

$$Ax - b = -b_\perp$$

$$A^*(Ax - b) = -A^* b_\perp \quad \text{① above}$$

$$A^*(Ax - b) = 0$$

hence

$$A^*A x = A^* b$$

for some x.