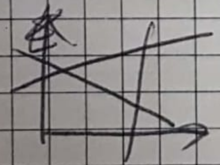
Least Squares Approximations & Least Norm Solution

$F\vec{\alpha} = \vec{y}$ 1. over-determined system $l > n$

①



$F^T F\alpha = F^T y \Rightarrow$

$\hat{\alpha} = (F^T F)^{-1} F^T y$

primal solution space

$F\vec{\alpha} = \vec{y}_0 + \vec{e} \Rightarrow \|F\vec{\alpha} - \vec{y}_0 - \vec{e}\| = \|\vec{e}\|$

② Under-determined system $l < n$

project $\alpha \in R^n$ into $\beta \in R^l$: $\alpha = F^T\beta$

, where $\beta \in R^l$

$FF^T\vec{\beta} = \vec{y} \Rightarrow \vec{\beta} = (FF^T)^{-1}\vec{y} \Rightarrow$

$\alpha = F^T(FF^T)^{-1}\vec{y}$

Least normal solution

Solving for Λ in the system of LE

$F\Lambda = B \quad F \in R^{m \times d} \quad \Lambda \in R^{d \times q}$

$B \in R^{m \times q}$

in the coleumn space (range) of F or in the row space (kernel) of F is equivalent to minimizing the scum of squarred errors (SSE) given by

$$SSE = trace\left((F\Lambda - B)^T (F\Lambda - B)\right)$$

Moreover the resultant solution $\hat{\Lambda}$ is unique with minimal-norm value in the sense that

$$\|\hat{\Lambda}\|_2^2 \leq \|\Lambda\|_2^2 \quad \text{for all feasible } \Lambda$$

Scanned by CamScanner

Multidimensional Regression

$f_1(x), \dots f(x_n)$ – numeric features

$$f(x,\alpha) = \sum_{j=1}^{n} \alpha_j f_j(x) \quad \alpha \in R^n$$

Matrix designation

$$\underset{l \times n}{F} = \begin{pmatrix} f_1(x) & \dots & f_n(x_1) \\ - & - & - \\ f_1(x_l) & \dots & f_n(x_l) \end{pmatrix} \quad \underset{l \times 1}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_l \end{pmatrix} \quad \underset{n \times l}{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Squared error functional

$$Q(\alpha, X^l) = \sum_{i=1}^{l} (f(x_i,\alpha) - y_i)^2 = \|F\alpha - y\|^2 \to \min_\alpha$$

$$\Downarrow$$

$$\frac{\partial Q(\alpha)}{\partial \alpha} = 2F^T(F\alpha - y) = 0$$

$$\Downarrow$$

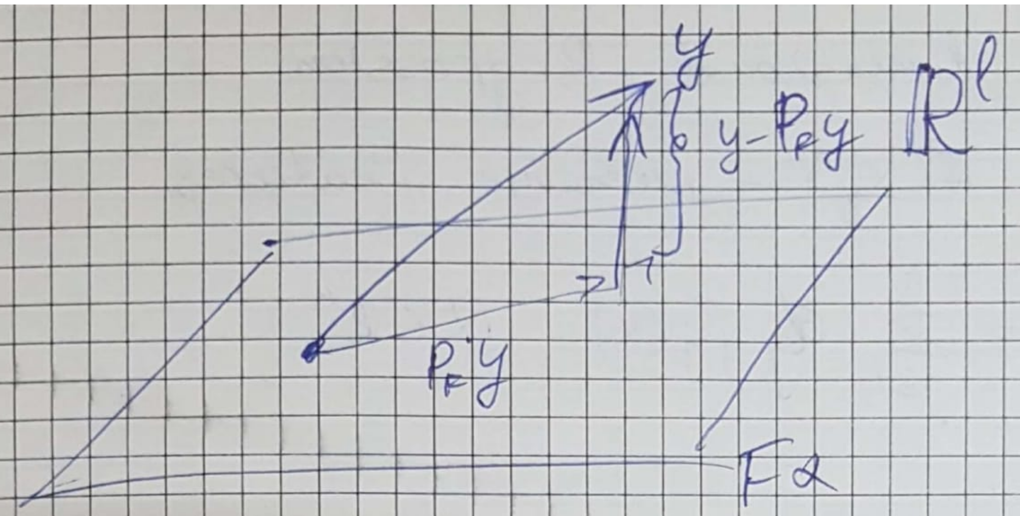
$$F^T F\alpha - F^T y = 0 \Rightarrow F^T F\alpha = F^T y$$

$\underset{n \times n}{F^T F}$ – covariation matrix of features set $f_1, \dots, f_n$ pseudo inverse matrix

The solution: $\alpha^* = (F^T F)^{-1} F^T y = F^+ y$

$Q(\alpha^*) = \|P_F y - y\|^2$ – functional

where $P_F = FF^+ = F(F^T F)^{-1} F^T$ – projection matrix



Давайте представим себе линейное многообразие натянутое на столбцы матрицы F - это все возможные векторы вида $F\alpha$

P_F - это проекция вектора "y" на линейное подпространство F

Singular Value Decomposition

Any Random matrix $\ell \times n$ can be decomposed as

$$F = VDU^T$$

Main Features of SVD:

1. $\ell \times n$ matrix $V = (v_1, \dots, v_n)$ - orthogonal, $V^T V = I_n$
columns v_j - eigen vectors of FF^T;

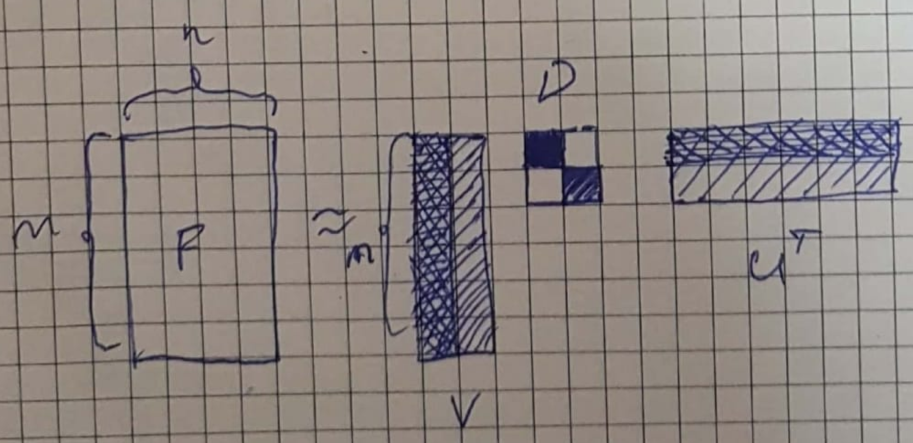
2. $n \times n$ matrix $U = (u_1, \dots, u_n)$ orthogonal $U^T U = I_n$
columns u_j - eigen vectors of $F^T F$

3. $n \times n$ D - диагональная $D = diag(\sqrt{\lambda_1}, \dots \sqrt{\lambda_n})$
$\lambda_j \geq 0$ - eigen values of $F^T F$ & $F F^T$

n - number of features
ℓ - number of objects

$$F_{[m\times n]} = V_{[m\times r]} D_{[n\times r]} \left(U_{[n\times r]}\right)^T = \sum_i d_i \vec{v}_i \circ \vec{u}_i^T$$

m - docs, n - terms



Solution LSM with the help of SVD

$$F^{+} = (UDV^{T}VDU^{T})^{-1}UDV^{T} = \quad \left\{ (AB)^{-1} = B^{-1}A^{-1} \right.$$

$$= UD^{-1}V^{T} = \sum_{j=1}^{n} \frac{1}{\sqrt{\lambda_j}} u_j v_j^{T}$$

$$\alpha^{*} = F^{+}y = UD^{-1}V^{T}y = \sum_{j=1}^{n} \frac{1}{\sqrt{\lambda_j}} u_j (v_j^{T}y)$$

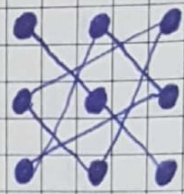
$$F\alpha^{*} = P_F y = (VDU^{T})UD^{-1}V^{T}y = VV^{T}y =$$

$$= \sum_{j=1}^{n} v_j (v_j^{T}y)$$

$$\|\alpha^{*}\|^{2} = \|D^{-1}V^{T}y\|^{2} = \sum_{j=1}^{n} \frac{1}{\lambda_j} (v_j^{T}y)^{2}$$

Singular Value Decomposition

$$A = \begin{bmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{SVD-?}$$



$$AA^T = \begin{bmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 2 \end{bmatrix}; \quad \det\left[AA^T - \lambda I\right] = -\lambda^3 + 10\lambda^2 - 16\lambda =$$

$$= -\lambda(\lambda^2 - 10\lambda + 16) = -\lambda(\lambda - 8)(\lambda - 2)$$

$$\boxed{A^T A = \begin{vmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & \sqrt{8} & 0 \\ \sqrt{8} & 6 & 2 \\ 0 & 2 & 2 \end{vmatrix}}$$

Eigenvalues of AA^T

$\lambda = 8, \quad \lambda = 2, \quad \lambda = 0$

Singular values $\sigma_1 = 2\sqrt{2}, \quad \sigma_2 = \sqrt{2} \quad \sigma_3 = 0$

$$\Sigma = \begin{vmatrix} \sqrt{8} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{vmatrix}$$

$$AA^T \vec{x} = \lambda_1 \vec{x}$$

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\begin{cases} \cancel{2}(x_1 + x_2 + x_3) = \overset{4}{\cancel{8}} x_1 \\ \cancel{2}x_1 + \overset{3}{\cancel{6}}x_2 + \cancel{2}x_3 = {}^4\cancel{8}x_2 \\ \cancel{2}(x_1 + x_2 + x_3) = \underset{4}{\cancel{8}} x_3 \end{cases}$$

$$\begin{cases} x_1 + x_2 + x_3 = 4x_1 \\ x_1 + 3x_2 + x_3 = 4x_2 \\ x_1 + x_2 + x_3 = 4x_3 \end{cases} \Rightarrow \quad \begin{aligned} & x_1 = x_3 = 1 \\ & x_1 - x_2 + x_3 = 0 \Rightarrow x_2 = 2 \\ & (1, 2, 1) \end{aligned}$$

$$\begin{cases} x_1 + x_2 + x_3 = x_1 \\ x_1 + 3x_2 + x_3 = x_2 \\ x_1 + x_2 + x_3 = x_3 \end{cases} \qquad \begin{cases} x_2 = -x_3 \\ x_1 + 2x_2 + x_3 = 0 \\ x_1 = -x_2 \end{cases} \qquad \begin{aligned} & x_1 = x_3 = -1 \\ & x_2 = -x_1 \\ & x_2 = 1 \end{aligned}$$

$$2x_1 + 2x_2 = 0$$

$$(-1, 1, -1)$$

Scanned by CamScanner

$\lambda = 0 \qquad (1, 0, -1)$

$(1, 2, 1)$

$(-1, 1, -1)$

$(1, 0, -1)$

$$P = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$A^T A = \begin{vmatrix} 2 & \sqrt{8} & 0 \\ \sqrt{8} & 6 & 2 \\ 0 & 2 & 2 \end{vmatrix} \qquad \lambda_1 = 8, \quad \lambda_2 = 2 \quad \lambda_3 = 0$$

$$q_1 = \left(\frac{1}{\sqrt{6}} \quad \frac{3}{\sqrt{12}} \quad \frac{1}{\sqrt{2}} \right)$$

$$q_2 = \left(\frac{1}{\sqrt{3}} \quad 0 \quad -\frac{2}{\sqrt{6}} \right)$$

$$q_3 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{2}, \frac{1}{2} \right)$$

$$Q = \begin{vmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{12}} & 0 & -\frac{1}{2} \\ \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{6}} & \frac{1}{2} \end{vmatrix}$$

$$A = P \Sigma Q^T$$

$$\vec{q}_i = \frac{1}{\sigma_i} A^T \cdot \vec{p}_i$$