Minimal Least Squares.
(MLS)

$$AX = Y,$$

X, Y are vectors.

But $X \in R^n$, $Y \in R^m$ and

$m > n$! A is (m, n) matrix.

More equations than unknowns, so we unlikely can solve the system. What to do?

Approximate!

$$\min_{X \in R^n} \| AX - Y \|^2$$

[redacted]

Or using sums:

$$\min_{x_1, x_2, \dots x_n} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A(i,j)\, x_j - Y_i \right)^2$$

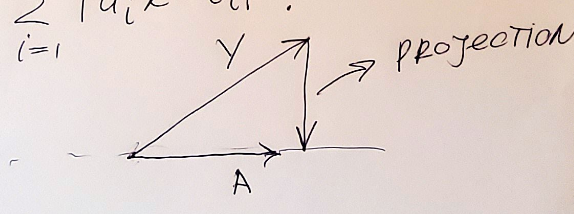
Example 1. We already had this problem for $n=1$:

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \qquad Y = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

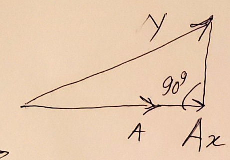
$$\downarrow$$

need to find at the minimum

$$\min_{x} \sum_{i=1}^{m} |a_i x - b_i|^2 :$$



We can solve this minimization problem using calculus, or geometrically:



$\vec{Y}$ is orthogonal to $\vec{Y} - \vec{Ax}$,

or $\langle Y, Y - Ax \rangle = 0$

$$\downarrow$$

$$x = \frac{\langle Y, A \rangle}{\langle A, A \rangle} \longrightarrow \text{optimal solution.}$$

In the general case $(n>1)$, the situation is very similar: instead of one vector, we have n vectors $\to$ columns of the matrix A:

$$A = \begin{bmatrix} V_1 & V_2 & \cdots & V_n \end{bmatrix}, \quad V_j \in \mathbb{R}^m.$$

And for the optimal solution $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ we get n orthogonality conditions:

$$\left\langle Y - \sum_{j=1}^{n} x_j V_j, V_k \right\rangle = 0 \quad \text{for all } 1 \le k \le n.$$

Putting those n orthogonality conditions into matrix-vector form, we get

$$A^T A \cdot X = A^T Y. \quad \text{(Verify!)}$$

Now, if the matrix $A^T A$ is nonsingular, which means that the vectors $(V_1, \dots, V_n)$ are linearly independent, then

$$\boxed{X = (A^T A)^{-1} A^T Y.}$$

And this is pretty much everything about MLS.

Example 2. Interpolation that is not exact:

$$f(x_i) = y_i, \quad 0 \le i \le n.$$

We can find out (always) the unique polynomial of degree n such that

$$p(x_i) = y_i, \quad 0 \le i \le n.$$

What about polynomial q of smaller than n degree? Not always. Istead we solve the minimization problem

$$\min_{b_0, \dots b_k} \sum_{j=0}^{n} \Big(\sum_{i=0}^{k} b_i x_j^i - y_j\Big)^2,$$

$k < n$.

IN A MATRIX FORM

$$A \begin{bmatrix} B_0 \\ \\ B_k \end{bmatrix} = \begin{bmatrix} Y_0 \\ \\ Y_n \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & X_0 & X_0^2 & \cdots & X_0^k \\ 1 & X_1 & X_1^2 & \cdots & X_1^k \\ \vdots & \vdots & & & \vdots \\ 1 & X_n & X_n^2 & \cdots & X_n^k \end{bmatrix}, \quad A \text{ is } (n+1, k+1).$$

IF $k \leq n$ THEN $A^T A$ IS NONSINGULAR (WHY?).

Example 3. There are many other ways to get/~~define~~ fine ~~a~~ approximation. But MLS is the simplest to compute!

For instance, instead of minimizing the error we could try to maximize the number of exactly satisfied equations (actually it is the minimization of the error in so called Hamming metric!

$$d(Y,Z) = \#\{i : Y_i \neq Z_i\}).$$

But this problem is NP-HARD.

THE SYSTEMS OF EQUATIONS WHERE IT IS NP-HARD TO MAXIMIZE THE NUMBER OF SATISFIED EQUATIONS COME FROM SO CALLED MAX-CUT PROBLEM: GIVEN AN UNDIRECTED GRAPH ON n VERTICES WITH m EDGES:

VERTICES $\to \{1, 2, \dots n\}$,

EDGES (PAIRS OF VERTICES) $\to \{(i_1, j_1), \dots (i_m, j_m)\}$,

$i_k < j_k$ FOR ALL $1 \leq k \leq m$.

ASSOCIATE WITH THE EDGE (i_k, j_k) 2 "FUNNY" EQUATIONS

$$x_{i_k} - x_{j_k} = 1$$

$$x_{i_k} - x_{j_k} = -1$$

ALL TOGETHER WE GET $2m$ EQUATIONS IN n UNKNOWNS.

MAXIMIZING THE NUMBER OF SATISFIED EQUATIONS IN THIS SYSTEM IS EQUIVALENT TO FINDING MAX-CUT!!!

(PERHAPS IT WAS COVERED IN YOUR 304 CLASS, ALSO RELATED TO 2-COLORING.)

APPLICATION OF SVD
(NOT COMMONLY DESCRIBED
IN TEXTBOOKS).

THE FOLLOWING two RESULTS
are used in computer vision.

Lemma 1. Let A, B be two (n,n)
PSD matrises.

$$A = U\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}U^T, \quad B = V\begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{pmatrix}V^T;$$

$UU^T = VV^T = I$; $a_1 \geq a_2 \geq \cdots a_n \geq 0$; $b_1 \geq b_2 \geq \cdots \geq b_n \geq 0$.

Then

$$\max_{KK^T=I} tr(KAK^TB) = \sum_{i=1}^{n} a_i b_i,$$

and optimal orthogonal $K = V \cdot U^T$.

Lemma 2.

Let C be (n,n) matrix, its Singular Value Decomposition

$$C = U\begin{pmatrix}\beta_1 & & 0\\ & \ddots & \\ 0 & & \beta_n\end{pmatrix}V:$$

$$UU^T = V^T V = I;\quad \beta_1 \geq \beta_2 \geq \dots \geq \beta_n \geq 0.$$

Then $\max\limits_{KK^T=I} tr(C\cdot K) = \beta_1 + \dots + \beta_n,$

and optimal $K = U^T V^T$.

(Why?)

Proof of Lemma 2.

$$tr(C\cdot K) = tr\left(\begin{pmatrix}\beta_1 & & 0\\ & \ddots & \\ 0 & & \beta_n\end{pmatrix}\cdot VK\cdot U\right) =$$

$$tr\left(\begin{pmatrix}\beta_1 & & 0\\ & \ddots & \\ 0 & & \beta_n\end{pmatrix}\cdot D\right),$$ where $D = V\cdot K\cdot U$.

Now, D is also orthogonal matrix:

$$DD^T = VK\underbrace{UU^T}_{=I}K^TV^T = I$$

Therefore

1. $|D(i,i)| \leq 1 \quad \forall\ 1 \leq i \leq n$

2. $D(i,i) = 1\ \forall 1 \leq i \leq n$ iff $D = I$.

Note that $tr(C \cdot K) = tr\left(\begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix} \cdot D\right) =$

$\sum_{i=1}^{n} \beta_i D(i,i)$ (Why?).

Thus $tr(C \cdot K) \leq \sum_{i=1}^{n} \beta_i$ and

$tr(CK) = \sum_{i=1}^{n} \beta_i$ if $VKU = I \Rightarrow K = V^TU^T$!!!

Example from Computer Vision.

Two images in 3D

$$[X_1|\cdots|X_m] =: \mathbb{X}$$

and $\qquad , \quad X_i, Y_i \in \mathbb{R}^3.$

$$[Y_1|\cdots|Y_m] =: \mathbb{Y}$$

We believe that image $\mathbb{Y}$ is the result of "physical move" of image $\mathbb{X}$:

$$Y_i = U X_i + S,$$

where U is the rotation matrix: $U^T U = I$ and $\det(U) = 1$; $S \in \mathbb{R}^3$.

THIS SYSTEM OF (NONLINEAR) EQUATIONS UNLIKELY HAS AN EXACT SOLUTION. SO, INSTEAD WE APPLY MINIMAL LEAST SQUARE (NONLINEAR) FITTING.

$$\min_{\substack{UU^T=I,\\ Det(U)=1,\\ S\in R^3}} \sum_{i=1}^{m} \| Y_i - U X_i - S \|_2^2 . \quad (***)$$

DEFINE $\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$, $\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$;

$\hat{Y}_i =: Y_i - \overline{Y}$, $\hat{X}_i = X_i - \overline{X}$, $1 \le i \le m$.

NOTE THAT

$$\sum_{i=1}^{m} \hat{Y}_i = \sum_{i=1}^{m} \hat{X}_i = 0. \; !!!$$

EXER. 1. PROVE THAT IF $U^T U = I$

THEN $$\sum_{i=1}^{m} \| Y_i - U X_i - S \|^2 =$$

$$= \sum_{i=1}^{m} \left(\| \hat{Y}_i \|_2^2 + \| \hat{X}_i \|_2^2 \right) +$$

$$m \| \bar{Y} - (U \bar{X} + S) \|_2^2 - tr\,(C \cdot U),$$

WHERE $$C = \sum_{i=1}^{m} \hat{X}_i (\hat{Y}_i)^T .$$

(We view here $\hat{X}_i$ and $\hat{Y}_i$ as

(3x1) matrices).

IT FOLLOWS FROM EXER. I THAT TO SOLVE OUR ORIGINAL MINIMIZATION PROBLEM (***), WE FIRST MINIMIZE

$$\min_{\substack{UU^T=I,\\ \det(U)=1}} -\mathrm{tr}(C\cdot U) = \mathrm{tr}(C\cdot \tilde{U});$$

AND COMPUTE $S=\bar{Y}-\tilde{U}\bar{X}$.

WITHOUT THE CONDITION $\mathrm{Det}(U)=1$ IT IS EXACTLY LEMMA 2, THE SOLUTION BOILS DOWN TO SVD FOR C:

$$\min_{UU^T=I} -\mathrm{tr}(C\cdot U) = -\max_{UU^T=I} \mathrm{tr}(C\cdot U).$$

To handle the condition $Det(U) = 1$ we need to use quaternions. Specifically, we need to use the quaternionic representation of (3×3) rotation matrices:

$$\mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \Longleftrightarrow \begin{pmatrix} x_1 & x_2+ix_3 \\ x_2-ix_3 & -x_1 \end{pmatrix}$$

(From 3D real vector to (2×2) complex hermitian matrix with zero trace.)

Rotations in 3D:

$$\begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} \Longleftrightarrow K = \begin{pmatrix} x_1 & x_2 + i x_3 \\ x_2 - i x_3 & -x_1 \end{pmatrix}$$

$$\downarrow$$

$$\tilde{K} = \begin{pmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{pmatrix} K \begin{pmatrix} \bar{z}_1 & \bar{z}_2 \\ -z_2 & z_1 \end{pmatrix},$$

where (the quaternion)

$$\begin{pmatrix} z_1 & -z_2 \\ z_2 & \bar{z}_1 \end{pmatrix} \longrightarrow z_1 = a + bi, \; z_2 = c + id,$$

$$a^2 + b^2 + c^2 + d^2$$

$$\tilde{K} = \begin{pmatrix} u & v + iw \\ v - iw & -u \end{pmatrix} \longrightarrow \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

SO, IF $\mathcal{U}$ IS (3×3) ROTATION MATRIX THEN THEN EXISTS $(a,b,c,d) \Rightarrow a^2+b^2+c^2+d^2=1$

SUCH THAT

$$\mathcal{U}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix},$$

AND THE ENTRIES OF $\mathcal{U}$ ARE QUADRATIC FORMS OF (a,b,c,d).

PUTTING ALL THIS TOGETHER, OUR ORIGINAL MINIMIZATION PROBLEM $(***)$

BOILS DOWN TO

~~min a^2+b^2~~

$$\min_{x_1^2+x_2^2+x_3^2+x_4^2} \sum_{1\le i,j\le 4} A(i,j)\, x_i x_j$$

FOR SOME SYMMETRIC (4×4)

REAL MATRIX: $d_1 \geq d_2 \geq d_3 \geq d_4$

$$A = U \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & d_3 & \\ 0 & & & d_4 \end{pmatrix} U^T \Longrightarrow$$

$\Longrightarrow$ SOLUTION $\min = d_4$,

OPTIMAL $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} U(1,4) \\ U(2,4) \\ U(3,4) \\ U(4,4) \end{pmatrix}$.