Recall: SVD for full-rank, tall matrix $A \in \mathbb{R}^{m \times n}$. Implication: $n$ # of singular values (non-zero)

\[ A = U \Sigma V^T \]

\[ \sum_{i=1}^{n} \sigma_i \begin{bmatrix} U_i^T \\ V_i \end{bmatrix} \]

Full SVD Reduced SVD

Sum of rank 1

Terminology: $U_i, V_i =$ singular vectors (≈ eigenvectors)

Q: In terms of SVD, what is $P$, the matrix which projects onto $\text{R}(A)$?

What sing. vals of $P$ & $A^*$?

\[ A^* = \begin{bmatrix} V \\ \Sigma^+ \\ U^* \end{bmatrix} \]

\[ \Sigma^+ = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{bmatrix} \]

Recall: \( \text{R}(A) = \text{Span} \left( U_1, \ldots, U_n \right) \) (since \( \text{rank}(A) = n \))

Orthobasis

\[ P = \sum_{i=1}^{n} U_i U_i^T = \begin{bmatrix} U \\ U^T \end{bmatrix} = \sum_{i=1}^{n} 1 \cdot U_i U_i^T \]
\[ P \text{ has singular values } \sigma_1 = \sigma_2 = \ldots = \sigma_n = 1. \]

\[ A^+ = V \Sigma^+ U^T = \sum_{i=1}^{n} \frac{1}{\sigma_i} V_i U_i^T \]

Sing values of \( A^+ \): \( \frac{1}{\sigma_n}, \frac{1}{\sigma_{n-1}}, \ldots, \frac{1}{\sigma_1} \), where \( \frac{1}{\sigma_n} \geq \frac{1}{\sigma_{n-1}} \)

**Accuracy of least squares assuming noisy linear model.**

\( y = Ax + z \)

\( y \), white noise (Gaussian/standard normal)

\( A \), known

\( z \), unknown \( \in \mathbb{R}^n \)

Fourier

**Def:** Given a random scalar \( q \), \( E_q \) is the expected value

\[ E_{z_i} = 0 \rightarrow \text{expected value of } z_i \text{ entries of noise} \]

\[ E_{z_i^2} = \text{noise level} = 1 \]

\( L \), S.D. (standard deviation)

**Note:** \((E_{z_i})^2 \neq E_{z_i^2} \)

**Lemma:** Let \( B \in \mathbb{R}^{n \times m} \).

\[ E \| B z \|^2 = \sum_{i=1}^{m} \sigma_i^2 (B) \]

\( \| B z \|^2 \), white noise

Function of \( B \)

\( E \), expected value

\[ E \| B z \|^2 = E \| B \Sigma U z \|^2 = E \| \Sigma V z \|^2 \]

\( B' \), SVD

\( B' \) does not change

\( \| B z \|^2 \) behaviour after rotation (\( V \))

Orthogonal matrix \( U \) preserves norm

\( \| B z \|^2 \)

Dimensional case (reduces to this)

\[ = E \sum_{i=1}^{m} \sigma_i^2 z_i^2 \]

\[ = \sum_{i=1}^{m} \sigma_i^2 E_{z_i^2} \]
Now, assume $m \geq n$, with $\text{rank}(A) = n$.

Let $\hat{x} = \arg \min \|A\hat{x} - y\|$ be the least squares estimate, $x^* \in \mathbb{R}^n$.

Goal: Determine $\|\|A\hat{x} - Ax\|_2^2$, also $\|Ax - Ax\|_2^2$.

~ How precise is my MRI image? On average, how large is the difference between my actual data and perceived image.

Recall: $\hat{x} = (A^TA)^{-1}A^Ty$ (from $A^TA\hat{x} = A^Ty$).

$= A^*(Ax + z)$

$= x + A^*z$

So, $\|Ax - x\|_2^2 = \|Ax + A^*z - x\|_2^2$

$= \|A^*z\|_2^2 = \sum_{i=1}^{n} \sigma_i(A)^2$ \text{ by lemma,}

$\sigma_i(A) = \frac{1}{\|P_i\|}$

\text{Rmk: If } A \text{ has small s.v.'s, it blows up the noise - badly conditioned } A$

Next, $\|A\hat{x} - Ax\|_2^2$

\text{worst case: noise lying in } \text{R}(A)

\text{best case: noise orthogonal to } \text{R}(A)$

$\tilde{z} = \text{proj}(A\hat{x} + z)$

$= \|Ax + A\tilde{z} - Ax\|_2^2$

$= \sum_{i=1}^{n} \sigma_i(P)^2$

$= n \cdot \dim(\text{proj}(\text{R}(A)))$
Elaboration on related research:

\[ y = Ax + z, x \in \mathbb{K} \]

Can we define \( \sigma_{k}(Ax) \) a non-restricted to \( \mathbb{K} \)?

→ projecting onto smaller-dimensional subspace, \( \mathbb{R}(A) \). → kills noise

Intuition, \( k = s = \text{subspace} \).

Is a subspace continuous? Yes, as it is defined by a set having close enough points (non-discrete)

Next class: FNN class w/ techniques on \( \mathbb{R}(A) \), \( N(A) \) “intuitions.”

Final - all-encompassing semi-proof test w/more time