SVD and rank

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}$$

The SVD is a <u>rank-revealing</u> factorization.

The rank of A is the number of positive singular values.

$m > n$

$$\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n & \vec{u}_{n+1} & \cdots & \vec{u}_m \\ | & | & & | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ \hline & & 0 & \end{bmatrix} \begin{bmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2^T & - \\ & \vdots & \\ - & \vec{v}_n^T & - \end{bmatrix}$$

A is <u>full rank</u> if $\text{rank}(A) = n$
(or $\text{rank}(A) = \min(m,n)$ more generally)

Then $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$

i.e., all the singular values are <u>positive</u>.

$$\begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | & | & & | \\ \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \\ | & & | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \\ \hline & & 0 & & & \end{bmatrix} \begin{bmatrix} - & \vec{v}_1^T & - \\ & \vdots & \\ - & \vec{v}_r^T & - \\ & \vdots & \\ - & \vec{v}_n^T & - \end{bmatrix}$$

A is <u>rank deficient</u> if $\text{rank}(A) < n$
(or $\text{rank}(A) < \min(m,n)$ more generally)

Then $\underbrace{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r}_{\text{positive singular values}} > \underbrace{\sigma_{r+1} = \cdots = \sigma_n}_{\text{zero singular values}} = 0$

In particular,

$\text{rank}(A) = r$

when there are exactly r positive singular values

Recall that another way to write the SVD is

$$A = \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T \quad , \quad A \in \mathbb{R}^{m\times n} \quad , \quad m \geq n$$

For A with $\text{rank}(A) = r$

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T + \sum_{i=r+1}^{n} 0 \cdot \vec{u}_i \vec{v}_i^T = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$$

If A is square and invertible (non-singular), $A \in \mathbb{R}^{n\times n}$

and

$$A = U \Sigma V^T$$

is the SVD of A, then

$$A^{-1} = V \Sigma^{-1} U^T,$$

where

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_n} \end{pmatrix}$$

if $A \in \mathbb{R}^{n\times n}$ invertible, note that $A^{-1} = A^+$
if A is $m\times n$, A^+ is $n\times m$

The notion of inverse can be extended to non-square and/or rank-deficient matrices through the SVD by defining the <u>pseudoinverse</u>, A^+, as

$$A^+ = V \Sigma^+ U^T, \quad [\Sigma^+]_{ij} = \begin{cases} [\Sigma]_{ji}^{-1} & \text{if } [\Sigma]_{ji}^{-1} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

So if

$$\underset{m\times n}{\Sigma} = \left[\begin{array}{c} \begin{matrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{matrix} \\ \hline 0 \end{array}\right], \quad \underset{n\times m}{\Sigma^+} = \left[\begin{array}{c|c} \begin{matrix} \frac{1}{\sigma_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_r} & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{matrix} & 0 \end{array}\right]$$

Projector

projector : square matrix that satisfies

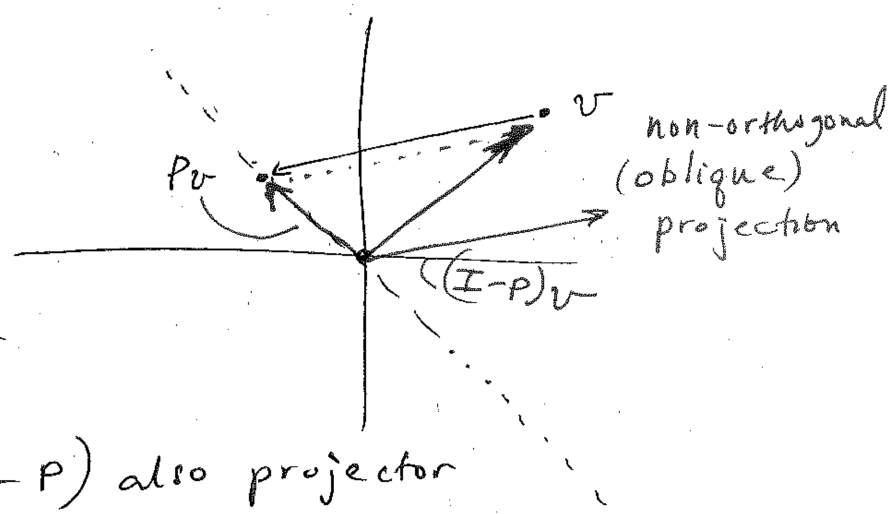
$$P^2 = P \qquad (\text{idempotent})$$

orthogonal projector + non-orthogonal or oblique projector

$v \in \text{Range}(P)$ projector does not change whatever is already in its range

$v = Px \quad \Rightarrow \quad Pv = P^2 x = Px = v$

$$P(Pv - v) = P^2 v - Pv = Pv - Pv = 0$$



Complementary Projector

P projector $\Rightarrow (I-P)$ also projector

$$(I-P)^2 = I - 2P + P^2 = I - 2P + P = I - P \checkmark$$

Orthogonal Projector

(not an orthogonal matrix)

range$(P) \perp$ range$(I-P)$

Theorem Projector P is orthogonal projector iff $P = P^T$

Proof show $Px \perp (I-P)y \quad \forall x, y \quad \Leftrightarrow \quad P = P^T$ ②

① $(Px)^T[(I-P)y] = x^T P^T (y - Py) = x^T P^T y - x^T P^T P y = 0$

Assume ②. Then $x^T P^T y - x^T P^T P y$

$= x^T P y - x^T P^2 y$

$= x^T P y - x^T P y \quad = 0 \Rightarrow$ ① ✓

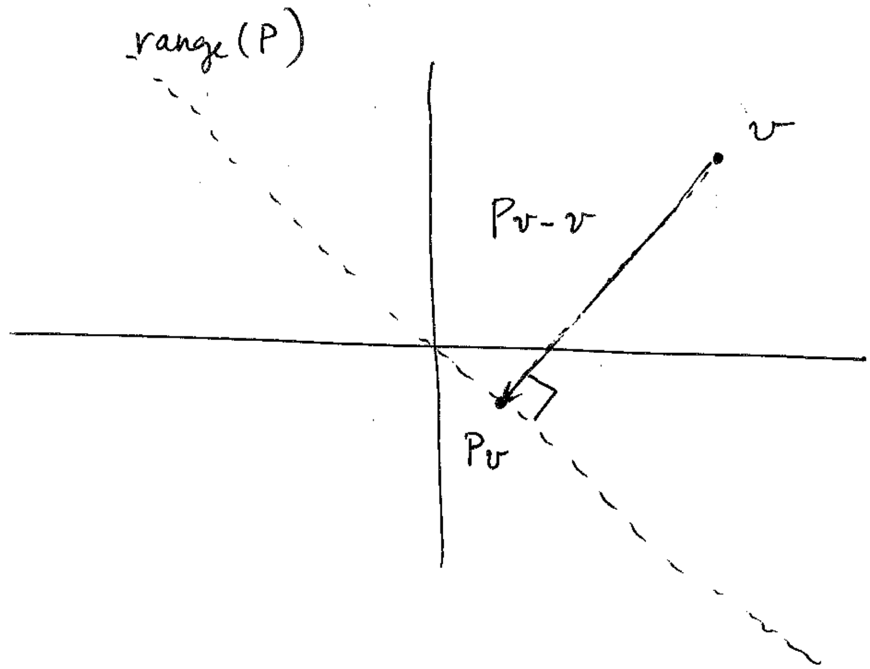
Assume ①. Then $x^T P^T y = x^T P^T P y \quad \forall x, y$

In particular $e_i^T P^T e_j = e_i^T P^T P e_j \quad \forall i, j \in \{1, \dots, n\}$

$\Rightarrow \quad P^T = P^T P$

taking transpose $P = P^T P$

$\Rightarrow P^T = P \Rightarrow$ ② ✓



orthogonal projection

Projection with an orthonormal basis.

orthogonal projector $P \quad (= P^T)$

$$P = U \Sigma U^T$$

and $\quad P^2 = U\Sigma U^T U \Sigma U^T = U \Sigma^2 U^T \qquad = PP^T$

$$\Rightarrow \quad \Sigma = \Sigma^2$$

singular values are all 0 or 1.

$(r = \text{rank}\, P \le n) \quad P = \sum_{i=1}^{r} \vec{u}_i \vec{u}_i^T = \hat{U}\hat{U}^T$, where $\hat{U} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{bmatrix}$

For any $\hat{Q}$ w/ orthonormal columns

$\hat{Q}\hat{Q}^T$ is an orthogonal projector onto columnspace of $\hat{Q}$.

Complementary projector $I - \hat{Q}\hat{Q}^T$ is also orthogonal projector

rank 1 $\quad P = (\vec{q}\vec{q}^T)$ $\qquad$ $\vec{q}$ unit vector

rank $n-1$ $\quad P^{\perp} = (I - \vec{q}\vec{q}^T)$

rank 1 or $\quad P = \dfrac{\vec{v}\vec{v}^T}{\vec{v}^T\vec{v}}$ $\qquad$ $\vec{v}$ arbitrary vector

normalize