Singular Value Decomposition (SVD)

If \( A \in \mathbb{C}^{m \times n} \), then there exists unitary matrices

\[
U \in \mathbb{C}^{m \times m} \quad \text{and} \quad V \in \mathbb{C}^{n \times n}
\]

such that

\[
A = U \Sigma V^H,
\]

where \( \Sigma \in \mathbb{C}^{m \times n} \) is the diagonal matrix of singular values.

In particular, if

\[
\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p
\]

are the non-zero singular values of \( A \) with

\[
\text{diag}(\Sigma) = [\sigma_1, \sigma_2, \sigma_3, \ldots]
\]

and

\[
U = [u_1, u_2, \ldots, u_m] \quad \text{and} \quad V = [v_1, v_2, \ldots, v_n],
\]

then

\[
\sigma_j u_j = Av_j \quad j = 1, 2, \ldots, p.
\]
Applications of the SVD

The SVD and Normal Matrices

Proposition. Let $A \in \mathbb{C}^{n \times n}$ be normal, and order the eigenvalues of $A$ as

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$ 

Then the singular values of $A$ are

$$\sigma_i = |\lambda_i|, \quad 1 \leq i \leq n.$$ 

Proof. By the Spectral Theorem for normal operators,

$$A = V \Lambda V^H,$$

where

$$V \in \mathbb{C}^{n \times n} \text{ is unitary and } V \in \mathbb{C}^{n \times n}.$$ 

For $1 \leq i \leq n$, choose $d_i \in \mathbb{C}$ for which

$$\bar{d}_i \lambda_i = |\lambda_i| \quad \text{and} \quad |d_i| = 1,$$

and let $D = \text{diag}(d_1, \ldots, d_n)$. Then $D$ is unitary, and

$$A = (VD)(D^H \Lambda)V^H \equiv U \Sigma V^H,$$

where $U = VD$ is unitary and

$$\Sigma = D^H \Lambda = \text{diag}(|\lambda_1|, \ldots, |\lambda_n|)$$

is diagonal with decreasing nonnegative diagonal entries. \qed
The SVD and Norms

The Frobenius and Euclidean operator norms of $A \in \mathbb{C}^{m \times n}$ are easily expressed in terms of the singular values of $A$:

$$\|A\|_F = \left( \sum_{i=1}^{n} \sigma_i^2 \right)^{\frac{1}{2}} = \left\| \begin{array}{c} \sigma_1 \\ \vdots \\ \sigma_n \end{array} \right\|_2$$

and

$$\|A\| = \sigma_1 = \sqrt{p(A^H A)} = \left\| \begin{array}{c} \sigma_1 \\ \vdots \\ \sigma_n \end{array} \right\|_\infty,$$

as follows from the unitary invariance of these norms.

There are no such simple expressions (in general) for these norms in terms of the eigenvalues of $A$ if $A$ is square (but not normal).
The SVD and Rank

The SVD is useful computationally for questions involving rank.

The rank of $A \in \mathbb{C}^{m \times n}$ is the number of nonzero singular values of $A$ since rank is invariant under pre- and post-multiplication by invertible matrices.

There are stable numerical algorithms for computing SVD (try \texttt{matlab}).

In the presence of round-off error, row-reduction to echelon form usually fails to find the rank of $A$ when its rank is $< \min(m, n)$.

For such a matrix, the computed SVD has the zero singular values computed to be on the order of machine $\epsilon$, and these are often identifiable as “numerical zeroes.”

For example, if the computed singular values of $A$ are $10^2, 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-15}, 10^{-15}, 10^{-16}$ with machine $\epsilon \approx 10^{-16}$, one can safely expect rank $(A) = 7$. 
The SVD and Polar Form

We now consider the matrix analogue of the polar form

\[ z = re^{i\theta}. \]

**Proposition**

Every \( A \in \mathbb{C}^{n \times n} \) may be written as

\[ A = PU, \]

where \( P \) is positive semi-definite Hermitian and \( U \) is unitary.

**Proof.** Let

\[ A = \Sigma V^H \]

be a SVD for \( A \), and write

\[ A = (U\Sigma U^H)(UV^H). \]

Then

\[ U\Sigma U^H \]

is positive semi-definite Hermitian and

\[ UV^H \]

is unitary. \( \square \)
Linear Least Squares Problems

If $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^{m}$, and consider the system

$$Ax = b.$$ 

This system may not be solvable, especially if $m > n$.

Instead solve

$$(*) \quad \inf_{x \in \mathbb{C}^{n}} ||Ax - b||^{2}_{2}.$$ 

This is called a least-squares problem since the square of the Euclidean norm is a sum of squares.

Set $\varphi(x) = ||Ax - b||^{2}$, then at the solution to $(*)$ $\nabla \varphi(x) = 0$, or equivalently

$$\varphi'(x; v) = 0 \quad \forall \ v \in \mathbb{C}^{n},$$

where

$$\varphi'(x; v) = \frac{d}{dt} \varphi(x + tv) \bigg|_{t=0}$$

is the directional derivative. If $y(t)$ is a differentiable curve in $\mathbb{C}^{m}$, then

$$\frac{d}{dt}||y(t)||^{2} = \langle y'(t), y(t) \rangle + \langle y(t), y'(t) \rangle = 2 \text{Re} \langle y(t), y'(t) \rangle.$$  

Taking $y(t) = A(x + tv) - b$, we obtain that

$$\nabla \varphi(x) = 0 \iff (\forall \ v \in \mathbb{C}^{n}) 2 \text{Re} \langle Ax - b, Av \rangle = 0 \iff A^{H}(Ax - b) = 0,$$

i.e.,

$$A^{H}Ax = A^{H}b.$$  

These are called the normal equations (they say $(Ax - b) \perp \mathcal{R}(A)$).
The Projection Theorem

Let $S \subset V$ be a subspace of the Euclidean space $V$.

(1) $V = S \oplus S^\perp$, i.e., given $v \in V$, $\exists$ unique $\bar{y} \in S$ and $\bar{z} \in S^\perp$ for which

$$v = \bar{y} + \bar{z}$$

(so $\bar{y} = Pv$, where $P$ is the orthogonal projection of $V$ onto $S$; also $\bar{z} = (I - P)v$ and $I - P$ is the orthogonal projection of $V$ onto $S^\perp$).

(2) Given $v \in V$, the $\bar{y}$ in (1) is the unique element of $S$ which satisfies

$$(\forall y \in S) \quad \langle v - \bar{y}, y \rangle = 0.$$  

(3) Given $v \in V$ let $\bar{y}$ be as in (1). Then $\hat{y} = \bar{y}$ if and only if $\hat{y}$ is the unique element of $S$ solving the minimization problem

$$\min_{y \in S} ||v - y||^2.$$
Proof of The Projection Theorem

(1) $V = S \oplus S^\perp$,

i.e., given $v \in V$, $\exists$ unique

$$\bar{y} \in S \quad \text{and} \quad \bar{z} \in S^\perp$$

for which

$$v = \bar{y} + \bar{z}$$

so

$$\bar{y} = Pv \quad \text{and} \quad \bar{z} = (I - P)v,$$

where $P$ is the orthogonal projection of $V$ onto $S$.

\textit{Proof}

Let $\{\psi_1, \ldots, \psi_r\}$ be an orthonormal basis of $S$.

Given $v \in V$, let

$$\bar{y} = \sum_{j=1}^{r} \langle v, \psi_j \rangle \psi_j \quad \text{and} \quad \bar{z} = v - \bar{y}.$$ 

Then $v = \bar{y} + \bar{z}$ and $\bar{y} \in S$.

For $1 \leq k \leq r$,

$$\langle \bar{z}, \psi_k \rangle = \langle v, \psi_k \rangle - \langle \bar{y}, \psi_k \rangle = \langle v, \psi_k \rangle - \langle v, \psi_k \rangle = 0,$$

so $\bar{z} \in S^\perp$.

Uniqueness follows from the fact that $S \cap S^\perp = \{0\}$. 
Proof of The Projection Theorem

(2) Given $v \in V$, the $\bar{y}$ in (1) is the unique element of $S$ which satisfies

$$\forall y \in S \quad \langle v - \bar{y}, y \rangle = 0.$$ 

Proof
Since $\bar{z} = v - \bar{y}$, this is just a restatement of $\bar{z} \in S^\perp$.

(3) Given $v \in V$ let $\bar{y}$ be as in (1). Then $\hat{y} = \bar{y}$ if and only if $\hat{y}$ is
the unique element of $S$ solving the minimization problem

$$\min_{y \in S} \|v - y\|^2.$$ 

Proof
For any $y \in S$,

$$v - y = \underbrace{\bar{y} - y}_{S} + \underbrace{\bar{z}}_{S^\perp},$$

so by the Pythagorean Theorem

$$(p \perp q \iff \|p \pm q\|^2 = \|p\|^2 + \|q\|^2),$$

$$\|v - y\|^2 = \|\bar{y} - y\|^2 + \|\bar{z}\|^2.$$ 

Therefore, $\|v - y\|^2$ is minimized iff $y = \bar{y}$, and $\|v - \bar{y}\|^2 = \|\bar{z}\|^2$. 
The Projection Theorem for Convex Sets

Let $X$ be a Hilbert space with $C \subset X$ closed and convex. Then there is a unique $y^0 \in C$ such that

$$\mathcal{P} \quad \|x - y^0\| \leq \|x - y\| \quad \forall \ y \in C.$$ 

Furthermore, $y_0$ satisfies $\mathcal{P}$ if and only if

$$\text{Re}(\langle x - y^0, y - y^0 \rangle) \leq 0 \quad \forall \ y \in C.$$ 

**Proof**

Let $\{y^i\} \subset C$ be such that

$$\|x - y^i\| \to \inf\{\|x - y\| : y \in C\} =: \delta.$$ 

By the parallelogram law

$$\|y^m - y^n\|^2 = 2\|x - y^m\|^2 + 2\|x - y^n\|^2 - 4\left\| x - \frac{y^n + y^m}{2} \right\|^2.$$ 

By convexity,

$$2^{-1}(y^n + y^m) \in C,$$ 

so

$$\|x - 2^{-1}(y^m + y^n)\| \geq \delta.$$ 

Therefore,

$$\|y^m - y^n\|^2 \leq 2\|y^m - x\|^2 + 2\|y^n - x\|^2 - 4\delta^2 \to 0.$$ 

Consequently, $\{y^n\}$ is Cauchy and so has a limit $y^0$ with

$$\|x - y^0\| = \delta.$$ 

Uniqueness follows by considering the sequence

$$y^{2n+1} = y^a \quad \text{and} \quad y^{2n} = y^b \quad n = 0, 1, \ldots$$ 

where $y^a, y^b \in C$ with $\|h^a - x\| = \|y^b - x\| = \delta$. Apply the above argument, to see that $y^a = y^b$. 
We now show that $y^0$ is the unique vector satisfying
$$\text{Re}(\langle x - y^0, y - y^0 \rangle) \leq 0$$
for all $y \in C$.
Suppose to the contrary that there is a vector $y^1$ such that
$$\text{Re}(\langle x - y^0, y^1 - y^0 \rangle) = \epsilon > 0.$$
Consider the vectors
$$y^\alpha = \alpha y^1 + (1 - \alpha)y^0 \in C \quad \text{for} \quad \alpha \in [0, 1].$$
Note that the function $\varphi : \mathbb{R} \to \mathbb{R}$ given by
$$\varphi(\alpha) = \| x - y^\alpha \|^2 = (1 - \alpha)^2\| x - y^0 \|^2 + 2\alpha(1 - \alpha)\text{Re}(\langle x - y^0, x - y^1 \rangle) + \alpha^2\| x - y^1 \|^2$$
is differentiable with
$$\varphi'(0) = -2\| x - y^0 \|^2 + 2\text{Re}(\langle x - y^0, x - y^1 \rangle)$$
$$= -2\text{Re}(\langle x - y^0, x - y^0 \rangle + \langle x - y^0, y^1 - x \rangle)$$
$$= -2\text{Re}(x - y^0, y^1 - y^0) = -2\epsilon < 0.$$ 
Hence, $\| x - y^\alpha \| < \| x - y^0 \|$ for all $\alpha > 0$ sufficiently small. This contradiction implies that $y^1$ does not exist.

Conversely, suppose that $y^0 \in C$ is such that
$$\text{Re}(\langle x - y^0, y - y^0 \rangle) \leq 0 \quad \forall \ y \in C.$$ 
Then for any $y \in C$ with $y \neq y^0$, we have
$$\| x - y \|^2 = \| (x - y^0) + (y^0 - y) \|^2$$
$$= \| x - y^0 \|^2 + 2\text{Re}(\langle x - y^0, y^0 - y \rangle) + \| y^0 - y \|^2$$
$$> \| x - y^0 \|^2.$$