Goals

- SVD-decomposition.
- Solving LLS with SVD-decomposition.
SVD Decomposition.

For any matrix $A \in \mathbb{R}^{m \times n}$ there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and a 'diagonal' matrix $\Sigma \in \mathbb{R}^{m \times n}$, i.e.,

$$
\Sigma = \begin{pmatrix}
\sigma_1 & & & & & \\
& \ddots & & & & \\
& & \sigma_r & & & \\
& & & 0 & & \\
& & & & \ddots & \\
& & & & & 0
\end{pmatrix}
$$

for $m \leq n$

with diagonal entries

$$
\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min\{m,n\}} = 0
$$

such that $A = U \Sigma V^T$
**SVD Decomposition.**

For any matrix $A \in \mathbb{R}^{m \times n}$ there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and a 'diagonal' matrix $\Sigma \in \mathbb{R}^{m \times n}$, i.e.,

$$
\Sigma = \begin{pmatrix}
\sigma_1 & & \\
& \ddots & \\
& & \sigma_r \\
0 & & \\
& & \ddots \\
0 & & & 0
\end{pmatrix}
$$

for $m \geq n$

with diagonal entries

$$
\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min\{m,n\}} = 0
$$

such that $A = U \Sigma V^T$
SVD Decomposition.

- The decomposition

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

is called **Singular Value Decomposition** (SVD). It is very important decomposition of a matrix and tells us a lot about its structure.

- It can be computed using the Matlab command `svd`.

- The diagonal entries \( \sigma_i \) of \( \Sigma \) are called the singular values of \( A \). The columns of \( U \) are called *left singular vectors* and the columns of \( V \) are called *right singular vectors*. 

The decomposition

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

is called **Singular Value Decomposition** (SVD). It is very important decomposition of a matrix and tells us a lot about its structure.

It can be computed using the Matlab command `svd`.

The diagonal entries \( \sigma_i \) of \( \Sigma \) are called the singular values of \( A \). The columns of \( U \) are called *left singular vectors* and the columns of \( V \) are called *right singular vectors*.

Using the orthogonality of \( V \) we can write it in the form

\[ AV = U \Sigma \]

We can interpret it as follows: there exists a special orthonormal set of vectors (i.e. the columns of \( V \)), that is mapped by the matrix \( A \) into an orthonormal set of vectors (i.e. the columns of \( U \)).
Applications of SVD Decomposition.

Given the SVD-Decomposition of $A$,

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

with

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min\{m,n\}} = 0$$

one may conclude the following:

- $\text{rank}(A) = r$,
- $\mathcal{R}(A) = \mathcal{R}([u_1, \ldots, u_r])$,
- $\mathcal{N}(A) = \mathcal{R}([v_{r+1}, \ldots, v_n])$,
- $\mathcal{R}(A^T) = \mathcal{R}([v_1, \ldots, v_r])$,
- $\mathcal{N}(A^T) = \mathcal{R}([u_{r+1}, \ldots, u_m])$. 
Moreover if we denote

\[ U_r = [u_1, \ldots, u_r], \quad \Sigma_r = \text{diag}(\sigma_1, \ldots, \sigma_r), \quad V_r = [v_1, \ldots, v_r], \]

then we have

\[ A = U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T \]

This is called the dyadic decomposition of \( A \), decomposes the matrix \( A \) of rank \( r \) into sum of \( r \) matrices of rank 1.
Applications of SVD Decomposition.

- The 2-norm and the Frobenius norm of $A$ can be easily computed from the SVD decomposition

\[
\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1
\]

\[
\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} = \sqrt{\sigma_1^2 + \cdots + \sigma_p^2}, \quad p = \min\{m, n\}.
\]

- From the SVD decomposition of $A$ it also follows that

\[
A^T A = V \Sigma^T \Sigma V^T \quad \text{and} \quad AA^T = U \Sigma \Sigma^T U^T.
\]

Thus, $\sigma_i^2, i = 1, \ldots, p$ are the eigenvalues of symmetric matrices $A^T A$ and $AA^T$ and $v_i$ and $u_i$ are the corresponding eigenvectors.
Applications of SVD Decomposition.

**Theorem**

Let the SVD of $A \in \mathbb{R}^{m \times n}$ be given by

$$A = U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

with $r = \text{rank}(A)$. If $k < r$

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T,$$

then

$$\min_{\text{rank}(D) = k} \| A - D \|_2 = \| A - A_k \|_2 = \sigma_{k+1},$$

and

$$\min_{\text{rank}(D) = k} \| A - D \|_F = \| A - A_k \|_F = \sqrt{\sum_{k+1}^{p} \sigma_i^2}, \quad p = \min \{m, n\}.$$
Consider the LLS

\[ \min_x \| Ax - b \|_2^2 \]

Let \( A = U \Sigma V^T \) be the SVD of \( A \in \mathbb{R}^{m \times n} \).

Using the orthogonality of \( U \) and \( V \) we have

\[
\| Ax - b \|_2^2 = \| U^T (AVV^T x - b) \|_2^2 = \| \sum V^T_x - U^T b \|_2^2 \\
= \sum_{i=1}^{r} (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^{m} (u_i^T b)^2.
\]
Thus, 

\[
\min_x \|Ax - b\|_2^2 = \sum_{i=1}^{r} (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^{m} (u_i^T b)^2.
\]
Solving LLS with SVD Decomposition.

Thus,

\[
\min_x \|Ax - b\|_2^2 = \sum_{i=1}^r (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^m (u_i^T b)^2.
\]

The solution is given

\[
z_i = \frac{u_i^T b}{\sigma_i}, \quad i = 1, \ldots, r,
\]

\[
z_i = \text{arbitrary}, \quad i = r + 1, \ldots, n.
\]
Solving LLS with SVD Decomposition.

Thus,

\[
\min_x \|Ax - b\|_2^2 = \sum_{i=1}^{r} (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^{m} (u_i^T b)^2.
\]

The solution is given

\[
z_i = \frac{u_i^T b}{\sigma_i}, \quad i = 1, \ldots, r,
\]
\[
z_i = \text{arbitrary}, \quad i = r + 1, \ldots, n.
\]

As a result

\[
\min_x \|Ax - b\|_2^2 = \sum_{i=r+1}^{m} (u_i^T b)^2.
\]
Recall that \( z = V^T x \). Since \( V \) is orthogonal, we find that
\[
\|x\|_2 = \|VV^T x\|_2 = \|V^T x\|_2 = \|z\|_2.
\]
All solutions of the linear least squares problem are given by \( z = V^T x \) with
\[
z_i = \frac{u_i^T b}{\sigma_i}, \quad i = 1, \ldots, r,
\]
\[
z_i = \text{arbitrary}, \quad i = r + 1, \ldots, n.
\]
Solving LLS with SVD Decomposition. Minimum norm solution

The minimum norm solution of the linear least squares problem is given by

\[ x^\dagger = Vz^\dagger, \]

where \( z^\dagger \in \mathbb{R}^n \) is the vector with entries

\[ z_i^\dagger = \frac{u_i^T b}{\sigma_i}, \quad i = 1, \ldots, r, \]
\[ z_i^\dagger = 0, \quad i = r + 1, \ldots, n. \]

The minimum norm solution is

\[ x^\dagger = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i \]
% compute the SVD:
[U,S,V] = svd(A);
s = diag(S);
% determine the effective rank r of A using singular values
r = 1;
while( r < size(A,2) & s(r+1) >= max(size(A))*eps*s(1) )
    r = r+1;
end
d = U'*b;
x = V* ( [d(1:r)./s(1:r); zeros(n-r,1) ] );
Conditioning of a Linear Least Squares Problem.

- Suppose that the data $b$ are

$$b = b_{ex} + \delta b,$$

where $\delta b$ represents the measurement error.

- The minimum norm solution of $\min \|Ax - (b_{ex} + \delta b)\|_2^2$ is

$$x^\dagger = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i = \sum_{i=1}^{r} \left( \frac{u_i^T b}{\sigma_i} + \frac{u_i^T \delta b}{\sigma_i} \right) v_i.$$
Conditioning of a Linear Least Squares Problem.

- Suppose that the data $b$ are

\[ b = b_{ex} + \delta b, \]

where $\delta b$ represents the measurement error.

- The minimum norm solution of $\min \|Ax - (b_{ex} + \delta b)\|_2^2$ is

\[ x^* = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i = \sum_{i=1}^{r} \left( \frac{u_i^T b}{\sigma_i} + \frac{u_i^T \delta b}{\sigma_i} \right) v_i. \]

- If a singular value $\sigma_i$ is small, then $\frac{u_i^T (\delta b)}{\sigma_i}$ could be large, even if $u_i^T (\delta b)$ is small. This shows that errors $\delta b$ in the data can be magnified by small singular values $\sigma_i$. 

Conditioning of a Linear Least Squares Problem.

% Compute A
\[ t = 10.^(0:-1:-10)'; \]
\[ A = [ \text{ones(size}(t)) \ t \ t.^2 \ t.^3 \ t.^4 \ t.^5]; \]
% compute SVD of A
\[ [U,S,V] = \text{svd}(A); \ \text{sigma} = \text{diag}(S); \]
% compute exact data
\[ xex = \text{ones}(6,1); \ bex = A \times xex; \]
for \( i = 1:10 \)
    % data perturbation
    \[ \text{deltab} = 10^(-i) \times (0.5-\text{rand(size}(bex))) \times bex; \]
    \[ b = bex + \text{deltab}; \]
    % solution of perturbed linear least squares problem
    \[ w = U' \times b; \]
    \[ x = V \times (w(1:6) ./ \text{sigma}); \]
    \[ \text{errx}(i+1) = \text{norm}(x - xex); \ \text{errb}(i+1) = \text{norm}(\text{deltab}); \]
end
\[ \text{loglog}(\text{errb}, \text{errx}, '\times'); \]
\[ \text{ylabel}('||x^{ex} - x||_2'); \ \text{xlabel}('||\delta b||_2') \]
Conditioning of a Linear Least Squares Problem.

- The singular values of $A$ in the above Matlab example are:

\[
\begin{align*}
\sigma_1 &\approx 3.4 & \sigma_4 &\approx 7.2 \times 10^{-4} \\
\sigma_2 &\approx 2.1 & \sigma_5 &\approx 6.6 \times 10^{-7} \\
\sigma_3 &\approx 8.2 \times 10^{-2} & \sigma_6 &\approx 5.5 \times 10^{-11}
\end{align*}
\]

- The error $\|x_{ex} - x\|_2$ for different values of $\|\delta b\|_2$ (loglog-scale):

![Graph showing the error $\|x_{ex} - x\|_2$ vs $\|\delta b\|_2$]

- We see that small perturbations $\delta b$ in the measurements can lead to large errors in the solution $x$ of the linear least squares problem if the singular values of $A$ are small.