EEM3L1: Numerical and Analytical Techniques

Lecture 5:
Singular Value Decomposition

SVD (1)
Motivation for SVD (1)

SVD = Singular Value Decomposition
Consider the system of linear equations $Ax = b$
Suppose $b$ is perturbed to $b + \delta b$
Solution becomes $x = A^{-1}b + A^{-1}\delta b$
The consequent change in $x$ is therefore $A^{-1}\delta b$
For what perturbation $\delta b$ will the error be biggest?
How big can the norm of the error be, in terms of $\|\delta b\|$?
The norm of the error, relative to $\|\delta b\|$ can be expressed in
terms of a number called the smallest singular value of $A$
Motivation for SVD (2)

Which direction $b$ must be perturbed in to give the biggest error?

If $\text{cond}(A)$ is large. How can we then find an accurate solution to $Ax = b$ ?

Both of these questions can also be addressed using Singular Value Decomposition

The remainder of the section of linear algebra will be taken up with Singular Value Decomposition (SVD)
Orthogonal Matrices revisited

Remember that an $m \times n$ matrix $U$ is called column orthogonal if $U^T U = I$, where $I$ is the identity matrix. In other words, the column vectors in $U$ are orthogonal to each other and each of them are of unit norm. If $n = m$ then $U$ is called orthogonal. In this case $UU^T = I$ also.
SVD of a Matrix

Let $A$ be an $m \times n$ matrix such that the number of rows $m$ is greater than or equal to the number of columns $n$. Then there exists:

(i) an $m \times n$ column orthogonal matrix $U$

(ii) an $n \times n$ diagonal matrix $S$, with positive or zero elements, and

(iii) an $n \times n$ orthogonal matrix $V$

such that: $A = USV^T$

This is the Singular Value Decomposition (SVD) of $A$
The Singular Values of $A$

Suppose $S = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$.
By convention it is assumed that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$
The values $\sigma_1, \sigma_2, \ldots, \sigma_n$ are called the singular values of $A$
SVD of a square matrix

The case where $A$ is an $n \times n$ square matrix is of particular interest.

In this case, the **Singular Value Decomposition** of $A$ is given

$$A = USV^T$$

Where $V$ and $U$ are **orthogonal** matrices
Example

Let \( A = \begin{bmatrix}
0.9501 & 0.8913 & 0.8214 & 0.9218 \\
0.2311 & 0.7621 & 0.4447 & 0.7382 \\
0.6068 & 0.4565 & 0.6154 & 0.1763 \\
0.4860 & 0.0185 & 0.7919 & 0.4057
\end{bmatrix} \)

In MATLAB, \( \gg \) \([U,S,V]=\text{svd}(A)\); returns the SVD of \( A \)

\[
U = \begin{bmatrix}
0.7301 & 0.1242 & 0.1899 & -0.6445 \\
0.4413 & 0.6334 & -0.3788 & 0.5104 \\
0.3809 & -0.3254 & 0.6577 & 0.5626 \\
0.3564 & -0.6910 & -0.6229 & 0.0871
\end{bmatrix}, \quad
V = \begin{bmatrix}
0.4903 & -0.4004 & 0.5191 & -0.5743 \\
0.4770 & 0.6433 & 0.4642 & 0.3783 \\
0.5362 & -0.5417 & -0.2770 & 0.5850 \\
0.4945 & 0.3638 & -0.6620 & -0.4299
\end{bmatrix}, \quad
S = \begin{bmatrix}
2.4479 & 0 & 0 & 0 \\
0 & 0.6716 & 0 & 0 \\
0 & 0 & 0.3646 & 0 \\
0 & 0 & 0 & 0.1927
\end{bmatrix}
\]
Singular values and Eigenvalues

The singular values of $A$ are not the same as its eigenvalues

```matlab
>> eig(A)
ans =
    2.3230
  0.0914+0.4586i
  0.0914-0.4586i
  0.2275
```

For any matrix $A$ the matrix $A^H A$ is normal with non-negative eigenvalues.

The singular values of $A$ are the square roots of the eigenvalues of $A^H A$
Calculating Inverses with SVD

Let $A$ be an $n \times n$ matrix.
Then $U$, $S$ and $V$ are also $n \times n$.
$U$ and $V$ are orthogonal, and so their inverses are equal to their transposes.
$S$ is diagonal, and so its inverse is the diagonal matrix whose elements are the inverses of the elements of $S$.

$$A^{-1} = V \left[ diag \left( \sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1} \right) \right] U^T$$
Calculating Inverses (contd)

If one of the $\sigma_i$s is zero, or so small that its value is dominated by round-off error, then there is a problem!
The more of the $\sigma_i$s that have this problem, the ‘more singular’ $A$ is.
SVD gives a way of determining how singular $A$ is.
The concept of ‘how singular’ $A$ is is linked with the condition number of $A$
The condition number of $A$ is the ratio of its largest singular value to its smallest singular value
The Null Space of $A$

Let $A$ be an $n \times n$ matrix
Consider the linear equations $Ax = b$, where $x$ and $b$ are vectors. The set of vectors $x$ such that $Ax = 0$ is a linear vector space, called the \textbf{null space} of $A$.
If $A$ is invertible, the null space of $A$ is the zero vector.
If $A$ is singular, the null space will contain non-zero vectors.
The dimension of the null space of $A$ is called the \textbf{nullity} of $A$. 
The Range of $A$

The set of vectors which are ‘targets’ for $A$, i.e. the set of all vectors $b$ for which there exists a vector $x$ such that $Ax=b$ is called the range of $A$

The range of $A$ is a linear vector space whose dimension is the rank of $A$

If $A$ is singular, then the rank of $A$ will be less than $n$

$n = \text{Rank}(A) + \text{Nullity}(A)$
SVD, Range and Null Space

Singular Valued Decomposition constructs orthonormal bases for the range and null space of a matrix.

The columns of $U$ which correspond to non-zero singular values of $A$ are an **orthonormal set of basis vectors for the range of $A$**.

The columns of $V$ which correspond to zero singular values form an **orthonormal basis for the null space of $A$**.
Solving linear equations with SVD

Consider a set of \textbf{homogeneous} equations $Ax = 0$. Any vector $x$ in the null space of $A$ is a solution. Hence any column of $V$ whose corresponding singular value is zero is a solution.

Now consider $Ax = b$ and $b \neq 0$.

A solution only exists if $b$ lies in the range of $A$. If so, then the set of equations does have a solution. In fact, it has infinitely many solutions because if $x$ is a solution and $y$ is in the null space of $A$, then $x + y$ is also a
Solving $Ax=b \neq 0$ using SVD

If we want a particular solution, then we might want to pick the solution $x$ with the smallest length $|x|^2$

Solution is

$$x = V \begin{bmatrix} \text{diag} (\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1}) \end{bmatrix} (U^T b)$$

where, for each singular value $\sigma_j$ such that $\sigma_j = 0$, $\sigma_j^{-1}$ is replaced by 0
Least Squares Estimation

If \( b \) is not in the range of \( A \) then there is no vector \( x \) such that \( Ax = b \). So

\[
x = V \left[ \text{diag} \left( \sigma_1^{-1}, \sigma_2^{-1}, ..., \sigma_n^{-1} \right) \right] (U^T b)
\]

cannot be used to obtain an exact solution.

**However, the vector returned will do the ‘closest possible job’ in the least squares sense.**

It will find the vector \( x \) which minimises \( R = \|Ax - b\| \)

\( R \) is called the **residual** of the solution.