Imperial College London

Maths for Signals and Systems

Linear Algebra in Engineering

Lectures 13 – 15, Tuesday 8th and Friday 11th November 2016

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING
IMPERIAL COLLEGE LONDON

Positive definite matrices

- **A symmetric or Hermitian matrix is positive definite if and only if (iff) all its eigenvalues are real and positive.**
- Therefore, the pivots are positive and the determinant is positive.
- However, positive determinant doesn't guarantee positive definiteness.

Example: Consider the matrix

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}$$

Eigenvalues are obtained from:

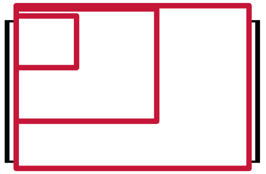
$$(5-\lambda)(3-\lambda) - 4 = 0 \Rightarrow \lambda^2 - 8\lambda + 11 = 0$$

$$\lambda_{1,2} = \frac{8 \pm \sqrt{64-44}}{2} = \frac{8 \pm \sqrt{20}}{2} = 4 \pm \sqrt{5}$$

The eigenvalues are positive and the matrix is symmetric, therefore, the matrix is positive definite.

Positive definite matrices cont.

- We are talking about symmetric matrices.
- We have various tests for positive definiteness. Consider the 2×2 case of a positive definite matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.
  - The eigenvalues are positive $\lambda_1 > 0, \lambda_2 > 0$.
  - The pivots are positive $a > 0$, $\frac{ac-b^2}{a} > 0$.
  - All determinates of leading ("north west") sub-matrices are positive $a > 0, ac - b^2 > 0$.



  - $x^T A\, x > 0$, x is any vector.
  - $x^T A\, x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2$. This is called **Quadratic Form**.

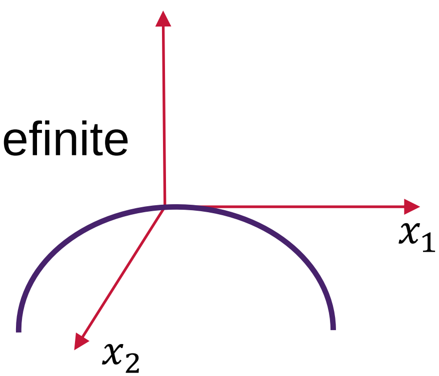
Positive semi-definite matrices

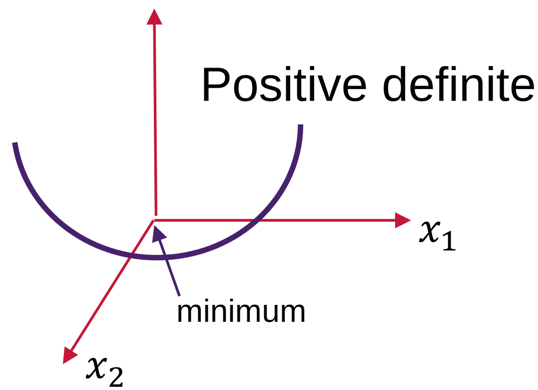
- **Example:** Consider the matrix $\begin{bmatrix} 2 & 6 \\ 6 & x \end{bmatrix}$
  - Which sufficiently large values of x makes the matrix positive definite? The answer is $x > 18$. (The determinant is $2x - 36 > 0 \Rightarrow x > 18$)
  - If $x = 18$ we obtain the matrix $\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$.
  - For $x = 18$ the matrix is **positive semi-definite**. The eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 20$. One of its eigenvalues is zero.
  - It has only one pivot since the matrix is singular. The pivots are 2 and 0.
  - Its quadratic form is $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 18x_2^2$.
  - In that case the matrix marginally failed the test.

Graph of quadratic form

- In mathematics, a **quadratic form** is a **homogeneous polynomial** of degree two in a number of variables. For example, the condition for positive-definiteness of a 2×2 matrix, $f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$, is a quadratic form in the variables x and y.

Not positive definite





- For the positive definite case we have:
  - Obviously, first derivatives must be zero at the minimum. This condition is not enough.
  - Second derivatives' matrix is positive definite, i.e., for $\begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix}$,

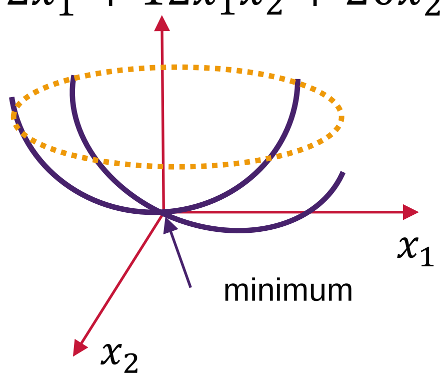
    we have $f_{x_1x_1} > 0, f_{x_1x_1}f_{x_2x_2} - 2f_{x_1x_2} > 0.$
  - Positive for a number turns into positive definite for a matrix.

Example 1

- **Example:**

$\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$, $\text{trace}(A) = 22 = \lambda_1 + \lambda_2$, $\det(A) = 4 = \lambda_1 \lambda_2$

- $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 20x_2^2$
- $f(x_1, x_1) = 2x_1^2 + 12x_1x_2 + 20x_2^2 = 2\left(x_1 \ + 3x_2\ \right)^2 + 2x_2^2.$



- A horizontal intersection could be $f(x_1, x_1) = 1$. It is an ellipse.
- Its quadratic form is $2\left(x_1 \ + 3x_2\ \right)^2 + 2x_2^2 = 1$.

Example 1 cont.

- Example:

$\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$, $\text{trace}(A) = 22 = \lambda_1 + \lambda_2$, $\det(A) = 4 = \lambda_1 \lambda_2$

- $f(x_1, x_2) = 2x_1^2 + 12x_1x_2 + 20x_2^2 = \mathbf{2}(x_1 + \mathbf{3}x_2)^2 + \mathbf{2}x_2^2$
- Note that computing the square form is effectively elimination

$A = \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \xrightarrow[(2)-3(1)]{} \begin{bmatrix} \mathbf{2} & 6 \\ 0 & \mathbf{2} \end{bmatrix} = U$ and $L = \begin{bmatrix} 1 & 0 \\ \mathbf{3} & 1 \end{bmatrix}$

- The **pivots** and the **multipliers** appear in the quadratic form when we compute the square.
- Pivots are the multipliers of the squared functions so positive pivots imply sum of squares and hence positive definiteness.

Example 2

- **Example:** Consider the matrix $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$
  - The leading ("north west") determinants are 2,3,4.
  - The pivots are 2, 3/2, 4/3.
  - The quadratic form is $\boldsymbol{x}^T A\, \boldsymbol{x} = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1\, x_2\, - 2x_2\, x_3$.
  - This can be written as:

$$\mathbf{2}\left(x_1 - \frac{1}{2}x_2\right)^2 + \frac{\mathbf{3}}{\mathbf{2}}\left(x_2 - \frac{2}{3}x_3\right)^2 + \frac{\mathbf{4}}{\mathbf{3}}x_3^2$$

  - The eigenvalues of A are $\lambda_1 = 2 - \sqrt{2}, \lambda_2 = 2, \lambda_3 = 2 + \sqrt{2}$
  - The matrix A is positive definite when $\boldsymbol{x}^T A\, \boldsymbol{x} > 0$.

Imperial College London

Positive definite matrices cont.

- If a matrix A is positive-definite, its inverse A^{-1} it also positive definite. This comes from the fact that the eigenvalues of the inverse of a matrix are equal to the inverses of the eigenvalues of the original matrix.

- If matrices A and B are positive definite, then their sum is positive definite. This comes from the fact $x^T(A+B)x = x^T Ax + x^T Bx > 0$. The same comment holds for positive semi-definiteness.

- Consider the matrix A of size $m \times n$, $m \neq n$ (rectangular, not square). In that case we are interested in the matrix $A^T A$ which is square.

- Is $A^T A$ positive definite?

The case of $A^T A$ and AA^T

- Is $A^T A$ positive definite?
- $x^T A^T Ax = (Ax)^T Ax = \|Ax\|^2$
- In order for $\|Ax\|^2 > 0$ for every $x \neq 0$, the null space of A must be zero.
- In case of A being a rectangular matrix of size $m \times n$ with $m > n$, the rank of A must be n.
- In case of A being a rectangular matrix of size $m \times n$ with $m < n$, the null space of A cannot be zero and therefore, $A^T A$ is not positive definite.
- Following the above analysis, it is straightforward to show that AA^T is positive definite if $m < n$ and the rank of A is m.

Similar matrices

- Consider two square matrices A and B.
- Suppose that for some invertible matrix M the relationship $B = M^{-1}AM$ holds. In that case we say that A and B are similar matrices.
- **Example:** Consider a matrix A which has a full set of eigenvectors. In that case $S^{-1}AS = \Lambda$. Based on the above A is similar to Λ.
- **Similar matrices have the same eigenvalues.**
- **Matrices with identical eigenvalues are not necessarily similar.**
- There are different families of matrices with the same eigenvalues.
- Consider the matrix A with eigenvalues λ and corresponding eigenvectors x and the matrix $B = M^{-1}AM$.

  We have $Ax = \lambda x \Rightarrow AMM^{-1}x = \lambda x \Rightarrow M^{-1}AMM^{-1}x = \lambda M^{-1}x$

  $$BM^{-1}x = \lambda M^{-1}x$$

  Therefore, λ is also an eigenvalue of B with corresponding eigenvector $M^{-1}x$.

Matrices with identical eigenvalues with some repeated

- Consider the families of matrices with repeated eigenvalues.
- **Example:** Lets take the 2×2 size matrices with eigenvalues $\lambda_1 = \lambda_2 = 4$.
  - The following two matrices

$$\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = 4I \text{ and } \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

    have eigenvalues $4,4$ but they belong to different families.
  - There are **<u>two</u>** families of matrices with eigenvalues $4,4$.
  - The matrix $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ has no "relatives". The only matrix similar to it, is itself.
  - The big family includes $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ and any matrix of the form $\begin{bmatrix} 4 & a \\ 0 & 4 \end{bmatrix}$, $a \neq 0$. These matrices are not diagonalizable since they only have one non-zero eigenvector.

Singular Value Decomposition (SVD)

- The so called **Singular Value Decomposition** (**SVD**) is one of the main highlights in Linear Algebra.

- Consider a matrix A of dimension $m \times n$ and rank r.

- I would like to diagonalize A. What I know so far is $A = S\Lambda S^{-1}$. This diagonalization has the following weaknesses:
  - A has to be square.
  - There are not always enough eigenvectors.
    - For example consider the matrix $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, $a \neq 0$. It only has the eigenvector $[x \quad 0]^T$.

- **Goal: I am looking for a type of decomposition which can be applied to any matrix.**

Singular Value Decomposition (SVD) cont.

- I am looking for a type of matrix factorization of the form $A = U\Sigma V^T$ where A is any real matrix A of dimension $m \times n$ and furthermore,
  - U is a unitary matrix $(U^T U = I)$ with columns u_i, of dimension $m \times m$.
  - Σ is an $m \times n$ rectangular matrix with non-negative real entries only along the main diagonal. The main diagonal is defined by the elements σ_{ij}, $i = j$.
  - V is a unitary matrix $(V^T V = I)$ with columns v_i, of dimension $n \times n$.
- U is, in general, different to V.
- The above type of decomposition is called **Singular Value Decomposition**.
- The non-zero elements of Σ are the so called **Singular Values** of matrix A. They are chosen to be **positive**.
- When A is a square invertible matrix then $A = S\Lambda S^{-1}$.
- When A is a symmetric matrix, the eigenvectors of S are orthonormal, so $A = Q\Lambda Q^T$.
- Therefore, for symmetric matrices SVD is effectively an eigenvector decomposition $U = Q = V$ and $\Lambda = \Sigma$.
- For complex matrices, **transpose** must be replaced with **conjugate transpose**.

Singular Value Decomposition (SVD) cont.

- From $A = U\Sigma V^T$, the following relationship hold:

$$AV = U\Sigma$$

- Do not forget that U and V are assumed to be unitary matrices and therefore,

$$U^TU = UU^T = V^TV = VV^T = I$$

- If I manage to write $A = U\Sigma V^T$, the matrix A^TA is decomposed as:

$$A^TA = V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T$$

- In the above expression $\Sigma^T\Sigma$ is a matrix of dimension $(n \times m) \times (m \times n) = n \times n$ (square matrix). From the form of the original Σ, you can easily deduct that $\Sigma^T\Sigma$ is a diagonal matrix. It has r non-zero elements across the diagonal. These are the squares of the singular values of A which are located along the main diagonal of the rectangular matrix Σ. $\Sigma^T\Sigma = \Sigma^2$ if the original matrix A is a square matrix.
- Please note the difference between the "diagonal" (square matrices) and the "main diagonal" (rectangular matrices).
- Therefore, the above expression is the eigenvector decomposition of A^TA as follows:

$$A^TA = V(\Sigma^T\Sigma)V^T$$

Singular Value Decomposition (SVD) cont.

- Similarly, the eigenvector decomposition of AA^T is:

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

- In the above expression $\Sigma\Sigma^T$ is a matrix of dimension $(m \times n) \times (n \times m) = m \times m$. Similarly to $\Sigma^T\Sigma$, it is a square matrix with r non-zero elements across the diagonal.

- Based on the properties stated in previous slides, the number and values of non-zero elements of matrices $\Sigma\Sigma^T$ and $\Sigma^T\Sigma$ are identical. Note that these two matrices have different dimensions if $m \neq n$. In that case one of them (the bigger one) has at least one zero element in its diagonal since they both have rank r which is $r \leq \min(m, n)$.

- From this and previous slides, we deduct that we can determine all the factors of SVD by the eigenvector decompositions of matrices $A^T A$ and AA^T.

Useful properties

- Let A be an $m \times n$ matrix and let B be an $n \times m$ matrix with $n \geq m$. Then the n eigenvalues of BA are the m eigenvalues of AB with the extra eigenvalues being 0. Therefore, the non-zero eigenvalues of AB and BA are identical.
- Therefore: Let A be an $m \times n$ matrix with $n \geq m$. Then the n eigenvalues of $A^T A$ are the m eigenvalues of AA^T with the extra eigenvalues being 0. Similar comments for $n \leq m$ are valid.
- Matrices A, $A^T A$ and AA^T have the same rank.
- Let A be an $m \times n$ matrix with $n \geq m$ and rank r. The matrix A has r **singular values**. Both $A^T A$ and AA^T have r non-zero eigenvalues which are the squares of the singular values of A. Furthermore:
  - $A^T A$ is of dimension $n \times n$. It has r eigenvectors $[v_1 \quad \cdots \quad v_r]$ associated with its r non-zero eigenvalues and $n - r$ eigenvectors associated with its $n - r$ zero eigenvalues.
  - AA^T is of dimension $m \times m$. It has r eigenvectors $[u_1 \quad \cdots \quad u_r]$ associated with its r non-zero eigenvalues and $m - r$ eigenvectors associated with its $m - r$ zero eigenvalues.

Singular Value Decomposition (SVD) cont.

- I can write $V = [v_1 \quad \dots \quad v_r \quad v_{r+1} \quad \dots \quad v_n]$ and $U = [u_1 \quad \dots \quad u_r \quad u_{r+1} \quad \dots \quad u_m]$.
- Matrices U and V have already been defined previously.
- Note that in the above matrices, I put first in the columns the eigenvectors of $A^T A$ and AA^T which correspond to non-zero eigenvalues.
- To take the above even further, I order the eigenvectors according to the magnitude of the associated eigenvalue.
- The eigenvector that corresponds to the maximum eigenvalue is placed in the first column and so on.
- This ordering is very helpful in various real life applications.

Singular Value Decomposition (SVD) cont.

- As already shown, from $A = U\Sigma V^T$ we obtain that $AV = U\Sigma$ or

$$A[v_1 \quad \dots \quad v_r \quad v_{r+1} \quad \dots \quad v_n] = [u_1 \quad \dots \quad u_r \quad u_{r+1} \quad \dots \quad u_m]\Sigma$$

- Therefore, we can break $AV = U\Sigma$ into a set of relationships of the form $Av_i = \sigma_i u_i$. Note that σ_i is a scalar and v_i and u_i vectors.

- For $i \leq r$ the relationship $AV = U\Sigma$ tells us that:
  - The vectors $v_1, v_2, \dots, v_r$ are in the row space of A. This is because from $AV = U\Sigma$ we have $U^T AVV^T = U^T U\Sigma V^T \Rightarrow U^T A = \Sigma V^T \Rightarrow v_i^T = \frac{1}{\sigma_i} u_i^T A,\ \sigma_i \neq 0$.

    Furthermore, since the v_i's associated with $\sigma_i \neq 0$ are orthonormal, they form a basis of the row space.
  - The vectors $u_1, u_2, \dots, u_r$ are in the column space of A. This observation comes directly from $u_i = \frac{1}{\sigma_i} Av_i,\ \sigma_i \neq 0$, i.e., u_is are linear combinations of columns of A. Furthermore, the u_is associated with $\sigma_i \neq 0$ are orthonormal. Thus, they form a basis of the column space.

Imperial College London

Singular Value Decomposition (SVD) cont.

- Based on the facts that:
  - $Av_i = \sigma_i u_i$,
  - v_i form an orthonormal basis of the row space of A,
  - u_i form an orthonormal basis of the column space of A, we conclude that:

  with SVD, an orthonormal basis of the row space, which is given by the columns of v, is mapped by matrix A to an orthonormal basis of the column space given by the columns of u. This comes from $AV = U\Sigma$.

- The $n - r$ additional v's which correspond to the zero eigenvalues of matrix A^TA are taken from the null space of A.

- The $m - r$ additional u's which correspond to the zero eigenvalues of matrix AA^T are taken from the left null space of A.

Examples of different Σ matrices

- We managed to find an orthonormal basis (V) of the row space and an orthonormal basis (U) of the column space that diagonalize the matrix A to Σ.

- In general, the basis of V is different to the basis of U.

- The SVD is written as:

$$A[v_1 \quad \dots \quad v_r \quad v_{r+1} \quad \dots \quad v_n] = [u_1 \quad \dots \quad u_r \quad u_{r+1} \quad \dots \quad u_m]\Sigma$$

- The form of matrix Σ depends on the dimensions m, n, r. It is of dimension $m \times n$. Its elements are chosen as:

$$\Sigma_{ij} = \begin{cases} \sqrt{\sigma_i^2} = \sigma_i & i = j, 1 \leq i, j \leq r \\ 0 & \text{otherwise} \end{cases}$$

  - σ_i^2 are the non-zero eigenvalues of $A^T A$ or AA^T.
  - $\sqrt{\sigma_i^2}$ are the non-zero singular values of A.

Examples of different Σ matrices cont.

- Example: $m = n = r = 3$.

$$\Sigma = \begin{bmatrix} \sqrt{\sigma_1^2} & 0 & 0 \\ 0 & \sqrt{\sigma_2^2} & 0 \\ 0 & 0 & \sqrt{\sigma_3^2} \end{bmatrix}$$

- Example: $m = 4, n = 3, r = 2$.

$$\Sigma = \begin{bmatrix} \sqrt{\sigma_1^2} & 0 & 0 \\ 0 & \sqrt{\sigma_2^2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Imperial College London

Truncated or Reduced Singular Value Decomposition

- In the expression for SVD we can reformulate the dimensions of all matrices involved by ignoring the eigenvectors which correspond to zero eigenvalues.
- In that case we have:

$$A\,[v_1 \quad \dots \quad v_r] = [u_1 \quad \dots \quad u_r]\begin{bmatrix}\sigma_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \sigma_r\end{bmatrix} \Rightarrow A = u_1\sigma_1 v_1^T + \cdots + u_r\sigma_r v_r^T$$

  where:

  - The dimension of A is $m \times n$.
  - The dimension of $[v_1 \quad \dots \quad v_r]$ is $n \times r$.
  - The dimension of $[u_1 \quad \dots \quad u_r]$ is $m \times r$.
  - The dimension of Σ is $r \times r$.
- The above formulation is called **Truncated or Reduced Singular Value Decomposition**.
- As seen, the Truncated SVD gives the splitting of A into a sum of r matrices, each of rank 1.
- In the case of a square, invertible matrix ($m = n = r$), the two decompositions are identical.

Singular Value Decomposition. Example 1.

- Example: $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$ and $A^T A = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$
- The eigenvalues of $A^T A$ are $\sigma_1^2 = 32$ and $\sigma_2^2 = 18$.
- The eigenvectors of $A^T A$ are $v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

$$A^T A = V \Sigma^2 V^T$$

- Similarly $AA^T = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$
- Therefore, the eigenvectors of AA^T are $u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $AA^T = U \Sigma^2 U^T$.
- **CAREFUL:** u_i's are chosen to satisfy the relationship $u_i = \frac{1}{\sigma_i} A v_i$, $i = 1,2$.
- Therefore, the SVD of $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$ is:

$$A = U \Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$$

Singular Value Decomposition. Example 2.

- Example: $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$ (singular) and $A^T A = \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} = \begin{bmatrix} 80 & 60 \\ 60 & 45 \end{bmatrix}$
- The eigenvalues of $A^T A$ are $\sigma_1^2 = 125$ and $\sigma_2^2 = 0$.
- The eigenvectors of $A^T A$ are $v_1 = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix}$

$$A^T A = V\Sigma^2 V^T$$

- Similarly $AA^T = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}\begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 25 & 50 \\ 50 & 100 \end{bmatrix}$
- u_1 is chosen to satisfy the relationship $u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{125}}\begin{bmatrix} 5 \\ 10 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.
- u_2 is chosen to be perpendicular to u_1. Note that the presence of u_2 and v_2 does not affect the calculations, since their elements are multiplied by zeros.
- Therefore, the SVD of $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$ is:

$$A = U\,\Sigma\,V^T = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 5\sqrt{5} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 4/5 & 3/5 \\ 4/5 & -3/5 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$$

Singular Value Decomposition. Example 2 cont.

- The SVD of $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$ is:

$$A = U\,\Sigma\,V^T = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 5\sqrt{5} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 \\ 4/5 & -3/5 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$$

- The truncated SVD is:

$$A = U\,\Sigma\,V^T = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} 5\sqrt{5} [4/5 \quad 3/5] = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$$

Singular Value Decomposition. Example 3.

- Example: $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. We see that $r = 2$. $A^T A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.
- The eigenvalues of $A^T A$ are $\sigma_1^2 = 3$ and $\sigma_2^2 = 1$ and $\sigma_3^2 = 0$ (obviously).
- The eigenvectors of $A^T A$ are $v_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$ and $v_3 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$.
- Similarly $AA^T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.
- u_1 is chosen to satisfy the relationship $u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- u_2 is chosen to satisfy the relationship $u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Note that the presence of v_3 does not affect the calculations, since its elements are multiplied by zeros.

Singular Value Decomposition. Example.

- Therefore, the SVD of $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

- The truncated SVD for this example is:

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

Pseudoinverse

- Suppose that A is a matrix of dimension $m \times n$ and rank r. The SVD of matrix A is given by:

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

- I define a matrix Σ^+ of dimension $n \times m$ as follows:

$$\Sigma_{ij}^+ = \begin{cases} 1/\sqrt{\sigma_i^2} & i = j, 1 \leq i,j \leq r \\ 0 & \text{otherwise} \end{cases}$$

- The matrix $A^+ = V\ \Sigma^+ U^T$ is called the **Pseudoinverse** of matrix A or the **Moore Penrose** inverse.
- $A^+A = V\ \Sigma^+ U^T\ U\Sigma V^T = V\ \Sigma^+ \Sigma V^T$.
- The matrix $\Sigma^+\Sigma$ is of dimension $n \times n$ (square) and has rank r. It is defined as follows:

$$(\Sigma^+\Sigma)_{ij} = \begin{cases} 1 & i = j, 1 \leq i,j \leq r \\ 0 & \text{otherwise} \end{cases}$$

- $AA^+ = U\Sigma V^T V\ \Sigma^+ U^T = U\Sigma\ \Sigma^+ U^T$.
- The matrix $\Sigma\ \Sigma^+$ is of dimension $m \times m$ and has rank r. It is defined as $\Sigma^+\Sigma$ above.

Pseudoinverse

- Note that $\Sigma^+\Sigma$ and $\Sigma\,\Sigma^+$ have different dimensions.

- Note that $\Sigma^+\Sigma$ and $\Sigma\,\Sigma^+$ look like identity matrices where the last $(n - r)$ or $(m - r)$ diagonal elements have been replaced by zeros.

- If $m \geq n$ and the rank of A is n then $\Sigma^+\Sigma = I_{n\times n}$. In that case
$$A^+A = V\,\Sigma^+U^T\,U\Sigma V^T = V\,\Sigma^+\Sigma V^T = V\,V^T = I_{n\times n}.$$
Therefore, A^+ is a left inverse matrix of A.

- If $m \leq n$ and the rank of A is m then $\Sigma\,\Sigma^+ = I_{m\times m}$. In that case $AA^+ = I_{m\times m}$. Therefore, A^+ is a right inverse matrix of A.

Pseudoinverse cont.

- As already proved, the relationship $Av_i = \sigma_i u_i$, which comes directly from $AV = U\Sigma$, maps a vector from the row space to the column space.

- Similarly, from $A^+ = V\,\Sigma^+ U^T$ we get $A^+ U = V\,\Sigma^+$ and therefore, $A^+ u_i = \frac{1}{\sigma_i} v_i$.
  Therefore, the multiplication of a vector from the column space with the pseudo inverse A^+, gives a the vector in the row space.

Imperial College London

Other types of matrix inverses

Consider a matrix A of dimension $m \times n$ and rank r. The following cases hold:

- $r = m = n$. In that case $AA^{-1} = I = A^{-1}A$. The matrix A has a **two-sided inverse** or simply an **inverse**.
- $r = n < m$ (more rows thank columns)
  - The matrix has full column rank (independent columns).
  - Null space = $\{0\}$.
  - 0 or 1 solutions to $Ax = b$.

  In that case $A^T A$ of dimension $n \times n$ is invertible and A has a **left inverse** only.

$$\underbrace{(A^T A)^{-1}A^T}_{n \times m} \underset{m \times n}{A} = I_{n \times n}$$

$$A_{left}^{-1}\ A = I_{n \times n}$$

- $r = m < n$ (more columns than rows)
  - The matrix has full row rank (independent rows).
  - There are $n - m$ free variables.
  - Left null space = $\{0\}$.
  - Infinite solutions to $Ax = b$.
  - In that case AA^T of dimension $m \times m$ is invertible and A has a **right inverse** only.

$$\underset{m\ x\ n}{A} \underbrace{A^T(AA^T)^{-1}}_{n\ x\ m} = I_{m \times m}$$

$$A\ A_{right}^{-1} = I_{m \times m}$$