Statistical and Mathematical Methods

Statistical and Mathematical Methods for Data Science
DS5003

Dr. Nasir Touheed

Linear Algebra Review

Basic

- Linear algebra provides a way of compactly representing and operating on sets of linear equations.
- For example, consider the following system of equations:

$$\begin{aligned} 4x_1 - 5x_2 &= -13 \\ -2x_1 + 3x_2 &= 9. \end{aligned}$$

- In matrix notation, we can write the system more compactly as

$$Ax = b$$

with

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} \; and \; b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

Notations

- By $A \in \mathbb{R}_{m \times n}$ we denote a matrix with m rows and n columns, where the entries of A are real numbers.
- By $x \in \mathbb{R}_n$, we denote a vector with n entries. By convention, an n-dimensional vector is often thought of as a matrix with n rows and 1 column, known as a column vector.
- The ith element of a vector x is denoted x_i:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

Notations

- We use the notation $a_{ij}(or A_{ij}, A_{i,j},\ etc)$ to denote the entry of A in the i^{th} row and jth column:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

- We denote the jth column of A by a_j or $A_{:,j}$ and the ith row of A by a_i^T or $A_{i,:}$:

$$A = \begin{bmatrix} | & | & & | \\ a_{11} & a_{12} & \dots & a_{1n} \\ | & | & & | \end{bmatrix} \qquad A = \begin{bmatrix} -- & a_1^T & -- \\ -- & a_1^T & -- \end{bmatrix}$$

Vector Space

- Let V be an arbitrary nonempty set of objects for which two operations are defined:
- addition and multiplication by numbers called scalars.
- addition implies that u and v in V an object u + v, called the sum of u and v.
- By scalar multiplication we mean a rule for associating with each scalar k and each object u in V an object ku, called the scalar multiple of u by k.
- If the following axioms are satisfied by all objects u, v, w in V and all scalars k and m, then we call V a vector space and we call the objects in V. vectors

Vector Space Axioms

- If u and v are objects in V, then $u + v \in V$.
- $u + v = v + u$
- $u + (v + w) = (u + v) + w$
- There exists zero vector, $0 + u = u + 0 = u \; \forall \; u \; \in \; V$.
- For each u in V, there is an object -u in V, called a negative of u, such that $u + (-u) = (-u) + u = 0$.
- If k is any scalar and u is any object in V, then ku is in V.
- $k(u + v) = ku + kv$
- $(k + m)u = ku + mu$
- $k(mu) = (km)(u)$
- $1u = u$

Example

- Given $u + v = (u_1 + v_1 + 1, u_2 + v_2 + 1), ku = (ku_1, ku_2)$
  a. Compute $u + v$ and ku for $u = (0, 4), v = (1, -3)$, and $k = 2$.
  b. Show that $(0, 0) \neq 0$.
  c. Show that $(-1, -1) = 0$.

SubSpaces

- If W is a nonempty set of vectors in a vector space V, then W is a subspace of V if and only if the following conditions are satisfied.
- If u and v are vectors in W, then u + v is in W.
- If k is a scalar and u is a vector in W, then ku is in W.

Example

- Check if following forms a subspace
- All matrices of the form

$$\begin{bmatrix} a & 1 \\ b & 1 \end{bmatrix}$$

- All matrices of the form

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

Row Echolen Form

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 3 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 3 & -4 \end{pmatrix} \times(-1)$$

$$\equiv$$

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 0 & 1 & -5 \end{pmatrix} \times(1) \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & -3 \end{pmatrix}$$

$$R_3 - 1 \cdot R_1 \to R_3 \qquad R_3 - (-1) \cdot R_2 \to R_3$$

Finding inverse of A (Gauss-Jordan)

- Find the inverse of the following matrix by Gauss-Jordan Algorithm

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 3 & -4 \end{bmatrix}$$

Finding inverse of A (Gauss-Jordan)

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 3 & -4 \end{pmatrix}^{(-1)} = ?$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 1 & 0 \\ 1 & 3 & -4 & 0 & 0 & 1 \end{array}\right) \times(-1) \underset{R_3 - 1\cdot R_1 \to R_3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 1 & 0 \\ 0 & 1 & -5 & -1 & 0 & 1 \end{array}\right) \times(-1) \underset{R_2/(-1) \to R_2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 0 \\ 0 & 1 & -5 & -1 & 0 & 1 \end{array}\right) \times(-1) \underset{R_3 - 1\cdot R_2 \to R_3}{\sim}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & -3 & -1 & 1 & 1 \end{array}\right) \times\left(\frac{-1}{3}\right) \underset{R_3/(-3) \to R_3}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{array}\right) \times(2) \underset{R_2 - (-2)\cdot R_3 \to R_2}{\sim} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{2}{3} & \frac{-5}{3} & \frac{-2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{array}\right) \times(-1) \underset{R_1 - 1\cdot R_3 \to R_1}{\sim}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{2}{3} & \frac{-5}{3} & \frac{-2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{array}\right) \times(-2) \underset{R_1 - 2\cdot R_2 \to R_1}{\sim} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-2}{3} & \frac{11}{3} & \frac{5}{3} \\ 0 & 1 & 0 & \frac{2}{3} & \frac{-5}{3} & \frac{-2}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{array}\right)$$

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 3 & -4 \end{pmatrix}^{(-1)} = \begin{pmatrix} \frac{-2}{3} & \frac{11}{3} & \frac{5}{3} \\ \frac{2}{3} & \frac{-5}{3} & \frac{-2}{3} \\ \frac{1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{pmatrix}$$

Row and Column Space of matrix A

- Let A be an m x n matrix
- The Column Space of matrix A is the vector space spanned by the column vectors of matrix A. i.e. all linear combinations of the column vectors.
- Since each column vector has m components, C(A) is a subspace of R^m
- The Row Space of matrix A is the vector space spanned by the row vectors of matrix A.
- Since each row vector has n components, R(A) is a subspace of R^n

Nullspace and Rank

- Let A be an m x n matrix
- The Nullspace of matrix A is the vector space spanned by all the vectors x which satisfy $Ax = 0$.
- Since the vector x has n components, N(A) is a subspace of R^n
- The Left Nullspace of matrix A, or the nullspace of A^T, is the vector space spanned by all the vectors y which satisfy $A^T y = 0$.
- Since the vector y has m components, $N(A^T)$ is a subspace of R^m.
- The common dimension of the row space and column space of a matrix A is called the rank of A and is denoted by rank(A)
- the dimension of the null space of A is called the nullity of A and is denoted by nullity(A).

- A is a 5×3 matrix.
- The rank of matrix A $= 2$.
- The dimension of C(A) $=$
- The dimension of R(A) $=$
- The C(A) is a subspace of R^{-}, so is __.
- The R(A) is a subspace of R^3, so is __.
- The dimension of N(A) is
- The dimension of $N(A^T)$=

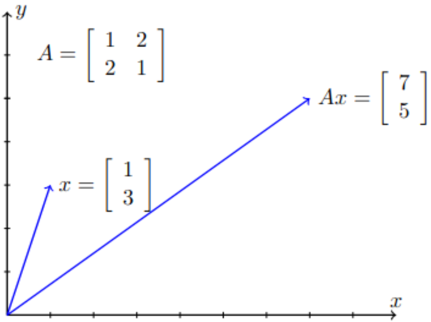
Example

- A is a 5×3 matrix.
- The rank of matrix A = 2.
- The dimension of C(A) = 2
- The dimension of R(A) = 2
- The C(A) is a subspace of R^5, so is$N(A^T)$.
- The R(A) is a subspace of R^3, so is N(A).
- The dimension of N(A) is 3-2=1
- The dimension of $N(A^T)$=5-2=3

Eigen Values and Eigen Vectors, Diagonalization

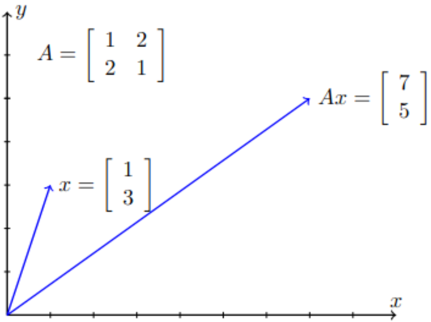
Matrix Operations

What happens when a matrix operates on a vector?



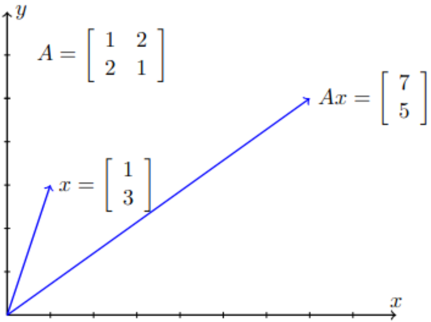
Matrix Operations

What happens when a matrix operates on a vector?
The vector gets transformed into a new vector (it strays from its path)



Matrix Operations

What happens when a matrix operates on a vector?
The vector gets transformed into a new vector (it strays from its path)
The vector may also get scaled (elongated or shortened) in the process.

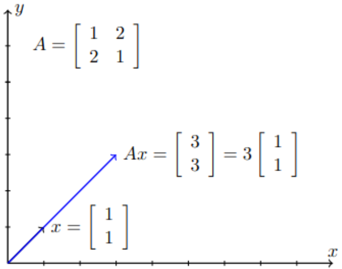


Eigen Values

- For a given square matrix A, there exist special vectors which refuse to stray from their path.
- These vectors are called eigenvectors.
- More formally,

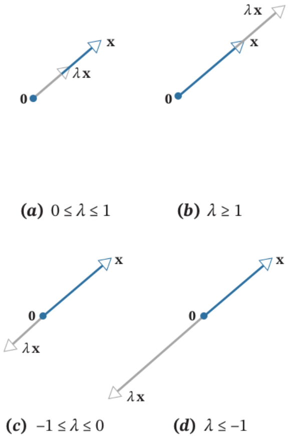
$$Ax = \lambda x$$

[direction remains the same]



Eigen Values

- In $\mathbb{R}^2$ or $\mathbb{R}^3$ multiplication by A maps each eigenvector x of A along the same line through the origin as x.
- Depending on the sign and magnitude of the eigenvalue λ corresponding to x,
- the operation $Ax = \lambda x$ compresses or stretches x by a factor of λ,
- Reversal of direction in the case where λ is negative



Eigen Values

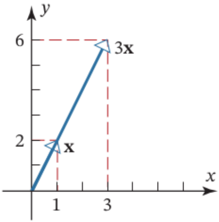
The vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

corresponding to the eigenvalue $\lambda = 3$, since

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$

Geometrically, multiplication by A has stretched the vector $\mathbf{x}$ by a factor of 3



Characteristic Equation

- The basic equation for eigenvectors and eigenvalues is

$$Ax = \lambda$$

- then

$$(A - \lambda I)x = 0$$

- So, the matrix $(A - \lambda I)$ has a nontrivial nulispace, and therefore must be singular.
- So,

$$det(A - \lambda I) = 0.$$

- So, if A is an eigenvalue of A then $det(A - \lambda I) = 0$.

Example

Eigen values of the matrix

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

Solving $(\lambda I - A)x = 0$ we have $\lambda = 3, -1$

$$\begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda = 3 \qquad \begin{bmatrix} 0 & 0 \\ -8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{aligned} x_1 &= \tfrac{1}{2}t \\ x_2 &= t \end{aligned}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

$\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ is a basis for the eigenspace corresponding to $\lambda = 3$

Example 2

Find bases for the eigenspaces of $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$

Solution:

- The characteristic equation of matrix A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$; thus, the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so there are two eigenspaces of A.

$$(\lambda I - A)\mathbf{x} = \mathbf{0} \Rightarrow \begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

If $\lambda = 2$, then (3) becomes $\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Example 2

Solving the system yield

$$x_1 = -s,\ x_2 = t,\ x_3 = s$$

Thus, the eigenvectors of A corresponding to $\lambda = 2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

The vectors $[-1\ 0\ 1]^T$ and $[0\ 1\ 0]^T$ are linearly independent and form a basis for the eigenspace corresponding to $\lambda = 2$.

Similarly, the eigenvectors of A corresponding to $\lambda = 1$ are the nonzero vectors of the form $\mathbf{x} = s\,[-2\ 1\ 1]^T$

Thus, $[-2\ 1\ 1]^T$ is a basis for the eigenspace corresponding to $\lambda = 1$.

Diagonalization

From the definition of the eigenvector $\boldsymbol{v}$ corresponding to the eigenvalue λ we have

$A\boldsymbol{v} = \lambda\boldsymbol{v}$ Then: $A\boldsymbol{v} - \lambda\boldsymbol{v} = (A - \lambda I) \cdot \boldsymbol{v} = 0$

Equation has a nonzero solution if and only if

$$\det(A - \lambda I) = 0$$

$$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -1 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & -1 & 3-\lambda \end{vmatrix}$$

$$= -\lambda^3 + 8\lambda^2 - 20\lambda + 16 = -(\lambda - 4)\cdot(\lambda^2 - 4\lambda + 4)$$

$$= -(\lambda - 4)\cdot(\lambda - 2)^2 = 0$$

$$\lambda_1 = 4 \quad \lambda_2 = 2$$

Diagonalization

For every λ we find its own vectors:

$\lambda_1 = 4$

$$A - \lambda_1 I = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -2 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

$\lambda_2 = 2$

$$A - \lambda_2 I = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$Av = \lambda v \qquad (A - \lambda I) \cdot v = 0$$

$$Av = \lambda v \qquad (A - \lambda I) \cdot v = 0$$

The solution set: $\{x_3 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\}$ $\{x_2 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\}$

Diagonalization

- The diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \lambda_3$

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

- The matrix with the Eigenvectors v_1, v_2, v_3 as its columns.

$$P = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

- hence

$$A = PDP^{-1}$$
$$A^2 = PDP^{-1}PDP^{-1}$$
$$A^2 = PD.DP^{-1}$$

The Singular Value Decomposition

- Let A be an $m \times n$ matrix with rank r.
- Then there exists an $m \times n$ matrix Σ for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq, \ldots \geq \sigma_r \, 0$,
- and there exist an m x m orthogonal matrix U and an n x n orthogonal matrix V such that

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

The Singular Value Decomposition

- Singular Value Decomposition is a decomposition $A = U * \Sigma * V^T$, where U and V are unitary matrices $(U * U^T = U^T * U = I$ and $V * V^T = V^T * V = I)$,
- Σ a diagonal matrix with non-negative entries.

$$
\begin{aligned}
A^T * A * V &= (U * \Sigma * V^T)^T * (U * \Sigma * V^T) * V \\
&= V * \Sigma^T * U^T * U * \Sigma * V^T * V \\
&= V * \Sigma^T * (U^T * U) * \Sigma * (V^T * V) \\
&= V * (\Sigma^T * \Sigma)
\end{aligned}
$$

- so $A^T * A * v_i = v_i * \sigma_i^2$ (where v_i is a column vector of V), which means v_i is an eigenvector of $A^T * A$ corresponding to an eigenvalue σ_i^2

The Singular Value Decomposition

$$A=\begin{bmatrix}2&2\\1&1\end{bmatrix}\Rightarrow A^TA=\begin{bmatrix}2&1\\2&1\end{bmatrix}\begin{bmatrix}2&2\\1&1\end{bmatrix}=\begin{bmatrix}5&5\\5&5\end{bmatrix}\text{ and } AA^T=\begin{bmatrix}2&2\\1&1\end{bmatrix}\begin{bmatrix}2&1\\2&1\end{bmatrix}=\begin{bmatrix}8&4\\4&2\end{bmatrix}$$

$$\det(A^TA-\lambda I)=(5-\lambda)^2-25=0\Rightarrow\lambda=0\text{ or }\lambda=10$$

$$\text{For }\lambda_1=0,\ \mathbf{x}_1=\begin{bmatrix}1\\-1\end{bmatrix}.\ \text{For }\lambda_2=10,\ \mathbf{x}_2=\begin{bmatrix}1\\1\end{bmatrix}.\Rightarrow V=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-1&1\end{bmatrix}$$

$$\det(AA^T-\lambda I)=(8-\lambda)(2-\lambda)-16=0\Rightarrow\lambda=0\text{ or }\lambda=10$$

$$\text{For }\lambda_1=0,\ \mathbf{x}_1=\begin{bmatrix}1\\-2\end{bmatrix}.\ \text{For }\lambda_2=10,\ \mathbf{x}_2=\begin{bmatrix}2\\1\end{bmatrix}.\Rightarrow U=\frac{1}{\sqrt{5}}\begin{bmatrix}1&2\\-2&1\end{bmatrix}$$

$$\text{Let }\Sigma=\begin{bmatrix}0&0\\0&\sqrt{10}\end{bmatrix}\Rightarrow U\Sigma V^{-1}=\frac{1}{\sqrt{5}}\begin{bmatrix}1&2\\-2&1\end{bmatrix}\begin{bmatrix}0&0\\0&\sqrt{10}\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-1&1\end{bmatrix}\right)^{-1}$$

$$=\begin{bmatrix}1&2\\-2&1\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}\begin{bmatrix}1&-1\\1&1\end{bmatrix}=\begin{bmatrix}0&2\\0&1\end{bmatrix}\begin{bmatrix}1&-1\\1&1\end{bmatrix}=\begin{bmatrix}2&2\\1&1\end{bmatrix}=A$$

The Singular Value Decomposition

- Find the singular value decomposition of

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$$