7. Symmetric Matrices and Quadratic Forms

7.1 Diagonalization of symmetric matrices

- Definition
  A symmetric matrix is a matrix ***A*** such that $\boldsymbol{A}^T = \boldsymbol{A}$.

- A symmetric matrix is necessarily square. Its main diagonal entries are arbitrary, but its other entries occur in pairs on opposite sides of the main diagonal.

- Example 1.

  Symmetric: $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} -1 & 6 & -2 \\ 6 & 2 & 0 \\ -2 & 0 & -3 \end{bmatrix}, \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}.$

  Nonsymmetric: $\begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} -1 & 6 & -2 \\ 6 & 2 & 0 \\ -3 & 0 & -3 \end{bmatrix}, \begin{bmatrix} a & d & e \\ -d & b & f \\ e & f & c \end{bmatrix}.$

- What is about eigenvectors of a symmetric matrix ?
- Theorem 1
  If $\boldsymbol{A}$ is symmetric, then any two distinct eigenvectors from different eigenspaces are orthogonal.

  Proof.
  Let $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ be eigenvectors that correspond to eigenvalues λ_1 and λ_2. To show that $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$, compute

$$\begin{aligned}\lambda_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 &= (\lambda_1 \boldsymbol{v}_1)^T \boldsymbol{v}_2 = (\boldsymbol{A} \boldsymbol{v}_1)^T \boldsymbol{v}_2 \\ &= (\boldsymbol{v}_1^T \boldsymbol{A}^T) \boldsymbol{v}_2 = \boldsymbol{v}_1^T (A \boldsymbol{v}_2) \\ &= \boldsymbol{v}_1^T (\lambda_2 \boldsymbol{v}_2) \\ &= \lambda_2 \boldsymbol{v}_1^T \boldsymbol{v}_2 = \lambda_2 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 .\end{aligned}$$

  Hence $(\lambda_1 - \lambda_2) \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$, but $\lambda_1 - \lambda_2 \neq 0$, so $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$.

- A matrix $\boldsymbol{A}$ is said to be orthogonally diagonalizable if there are an orthogonal matrix $\boldsymbol{P}$ (with $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$) and a diagonal matrix $\boldsymbol{D}$ such that

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}. \quad (1)$$

- To orthogonally diagonalize an $n \times n$ matrix, we must find n linearly independent and orthonormal eigenvectors.
- Moreover, if $\boldsymbol{A}$ is orthogonally diagonalizable as the above equation, then

$$\boldsymbol{A}^T = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T)^T = \boldsymbol{P}^{TT}\boldsymbol{D}^T\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T = \boldsymbol{A}.$$

- Theorem 2.
  An $n \times n$ matrix $\boldsymbol{A}$ is orthogonally diagonalizable if and only if $\boldsymbol{A}$ is a symmetric matrix.
  Proof.
  The proof is much harder and is omitted.
  Easy orthogonally diagonalizable $\Rightarrow$ symmetric
  Hard orthogonally diagonalizable $\Leftarrow$ symmetric

- Example 3.
  Orthogonally diagonalize the matrix $\boldsymbol{A} = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.
  Solution.

$$\lambda_1 = 7,\ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix};\ \lambda_2 = -2,\ \boldsymbol{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}.$$

Although $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are only linearly independent, not orthogonal; but they can be become orthogonal by the the Gram-Schmidt process.

$$\boldsymbol{v}_1 \Rightarrow \boldsymbol{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix},\ \boldsymbol{v}_2 \Rightarrow \boldsymbol{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix},\ \boldsymbol{v}_3 \Rightarrow \boldsymbol{u}_3 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}.$$

Then $\boldsymbol{P} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}$ and $\boldsymbol{D} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.

The spectral theorem

- The set of eigenvalues of a matrix $\boldsymbol{A}$ is sometimes called the *spectrum* of $\boldsymbol{A}$, and the following description of the eigenvalues is called a *spectral theorem*.
- Theorem 3. (*The spectral theorem for symmetric matrices*)
  An $n \times n$ symmetric matrix $\boldsymbol{A}$ has the following properties:
  a. $\boldsymbol{A}$ has n real eigenvalues, counting multiplicities.
  b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.
  c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.
  d. $\boldsymbol{A}$ is orthogonally diagonalizable.

Spectral decomposition

- Support $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$, where the columns of $\boldsymbol{P}$ are orthonormal eigenvectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ of $\boldsymbol{A}$ and the corresponding eigenvalues $\lambda_1, \ldots \lambda_n$ are in the diagonal matrix $\boldsymbol{D}$. Then, since $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$,

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^T = [\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n]\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & .. & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}\begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix}$$

$$= [\lambda_1\boldsymbol{u}_1 \ \cdots \ \lambda_n\boldsymbol{u}_n]\begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix}.$$

Ch.2, p.24, Theorem 10

Using the column-row expansion of a product, we can write

$$\boldsymbol{A} = \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^T + \lambda_2\boldsymbol{u}_2\boldsymbol{u}_2^T + \cdots + \lambda_n\boldsymbol{u}_n\boldsymbol{u}_n^T. \qquad (2)$$

This representation of $\boldsymbol{A}$ is called a spectral decomposition of $\boldsymbol{A}$. Each term in the above equation is a $n\times n$ matrix of rank 1. Every column of $\lambda_i\boldsymbol{u}_i\boldsymbol{u}_i^T$ is a multiple of $\boldsymbol{u}_i$. Furthermore, each matrix $\boldsymbol{u}_j\boldsymbol{u}_j^T$ is a projection matrix in the sense that for each $\boldsymbol{x}$ in $\boldsymbol{R}^n$, the vector $(\boldsymbol{u}_j\boldsymbol{u}_j^T)\,\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{x}$ onto the subspace spanned by $\boldsymbol{u}_j$. (Ch.6, p.18, Theorem 10)

- Example 4.
Construct a spectral decomposition of the matrix $\boldsymbol{A}$ that has the orthogonal diagonalization

$$\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Solution.
Denote the columns of $\boldsymbol{P}$ by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Then

$$\boldsymbol{A} = 8\,\boldsymbol{u}_1\,\boldsymbol{u}_1^T + 3\,\boldsymbol{u}_2\,\boldsymbol{u}_2^T$$

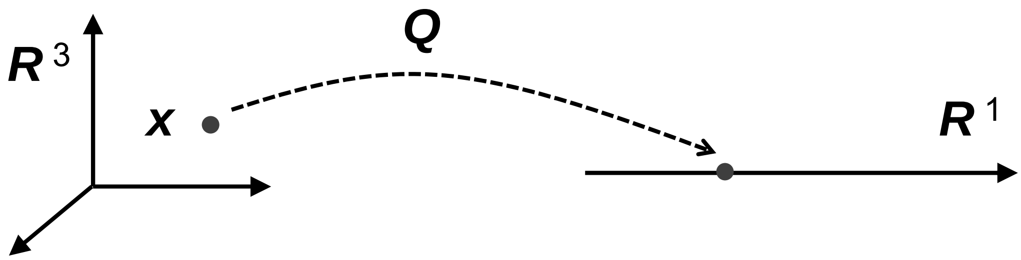
$$= 8\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} + 3\begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

$$= 8\begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix} + 3\begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}.$$

- Exercises of Section 7.1.

7.2 Quadratic forms

- A quadratic form on $\boldsymbol{R}^n$ is a function $\boldsymbol{Q}$ defined on $\boldsymbol{R}^n$ whose value at a vector $\boldsymbol{x}$ in $\boldsymbol{R}^n$ can be computed by an expression of the form $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A}$ is an nxn symmetric matrix. The matrix $\boldsymbol{A}$ is called the matrix of the quadratic form.



- The simplest example of a nonzero quadratic form is

$$\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{I}\,\boldsymbol{x} = ||\boldsymbol{x}||^2.$$

- Example 1.

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 + 2x_2 & -2x_1 + 7x_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^2 - 2x_1x_2 - 2x_1x_2 + 7x_2^2$$

$$= 3x_1^2 - 4x_1x_2 + 7x_2^2.$$

- Example 2. For $\boldsymbol{x}$ in $\boldsymbol{R}^3$,

$$5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3 =$$

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Change of variable in a quadratic form

- If $\boldsymbol{x}$ is a variable vector in $\boldsymbol{R}^n$, then a change of variable is an equation of the form
  $$\boldsymbol{x} = P\boldsymbol{y} \quad \text{or} \quad \boldsymbol{y} = P^{-1}\boldsymbol{x},$$
  where $\boldsymbol{P}$ is an invertible matrix and $\boldsymbol{y}$ is a new variable vector in $\boldsymbol{R}^n$.

- The $\boldsymbol{y}$ is the coordinate vector of $\boldsymbol{x}$ relative to the basis of $\boldsymbol{R}^n$ determined by the columns of $\boldsymbol{P}$. $[\boldsymbol{x}]_E = \boldsymbol{P}_\beta\,[\boldsymbol{x}]_\beta$

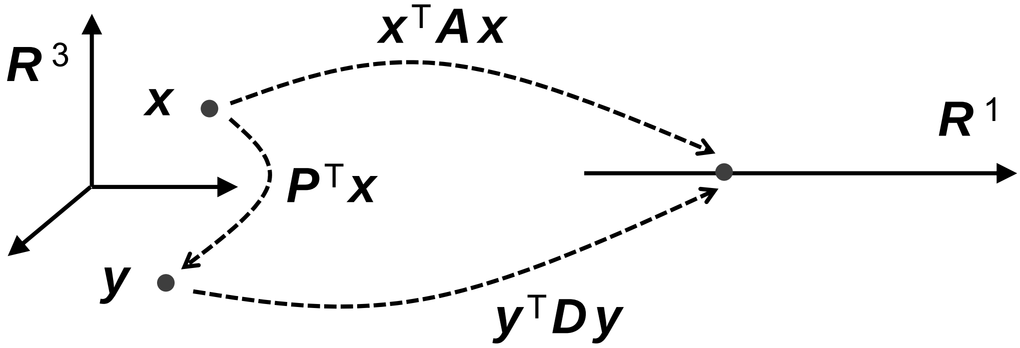
- If the change of variable is substituted into a quadratic form $\boldsymbol{x}^\top A\boldsymbol{x}$, then
  $$\boldsymbol{x}^\top A\boldsymbol{x} = (P\boldsymbol{y})^\top A(P\boldsymbol{y}) = \boldsymbol{y}^\top P^\top AP\boldsymbol{y} = \boldsymbol{y}^\top(P^\top AP)\boldsymbol{y}.$$
  Since $\boldsymbol{A}$ is orthogonally diagonalizable (Theorem 2), there is a orthonormal matrix $\boldsymbol{P}$ such that $P^\top AP$ is a diagonal matrix $\boldsymbol{D}$ and thus, the quadratic form becomes $\boldsymbol{y}^\top D\boldsymbol{y}$.

- Example 4.
  Make a change of variable that transforms a quadratic form with $A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$ into a quadratic form with no cross-product term.

  Solution.

  Step 1. Find eigenvalues and eigenvectors of $\boldsymbol{A}$
  $$\lambda_1 = 3,\ \mathbf{e}_1 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix};\quad \lambda_2 = -7,\ \mathbf{e}_2 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$
  Step 2. Construct $\boldsymbol{P}$ and $\boldsymbol{D}$,
  $$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix},\quad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$
  Step 3. Transform quadratic form with no cross-product term
  $$\boldsymbol{y}^T D\boldsymbol{y} = 3y_1^2 - 7y_2^2.$$

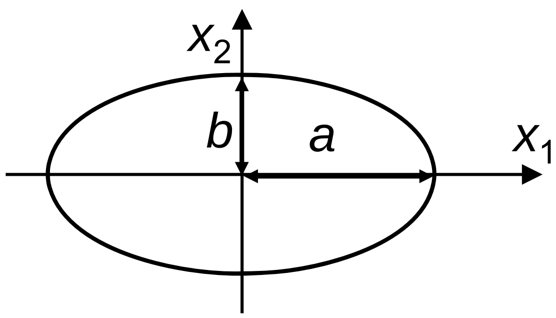
- **Theorem 4. (The principal axes theorem)**
  Let $\boldsymbol{A}$ be an $n\times n$ symmetric matrix. Then there is an orthogonal change of variable, $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$, that transforms the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ into a quadratic form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{y}$ with no cross-product term.

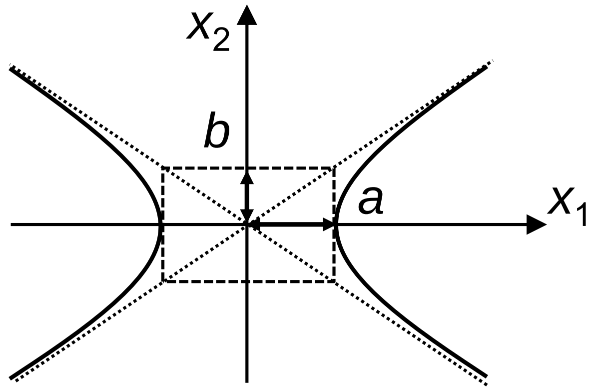


- The columns of $\boldsymbol{P}$ in Theorem 4 are called the principal axes of the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$.

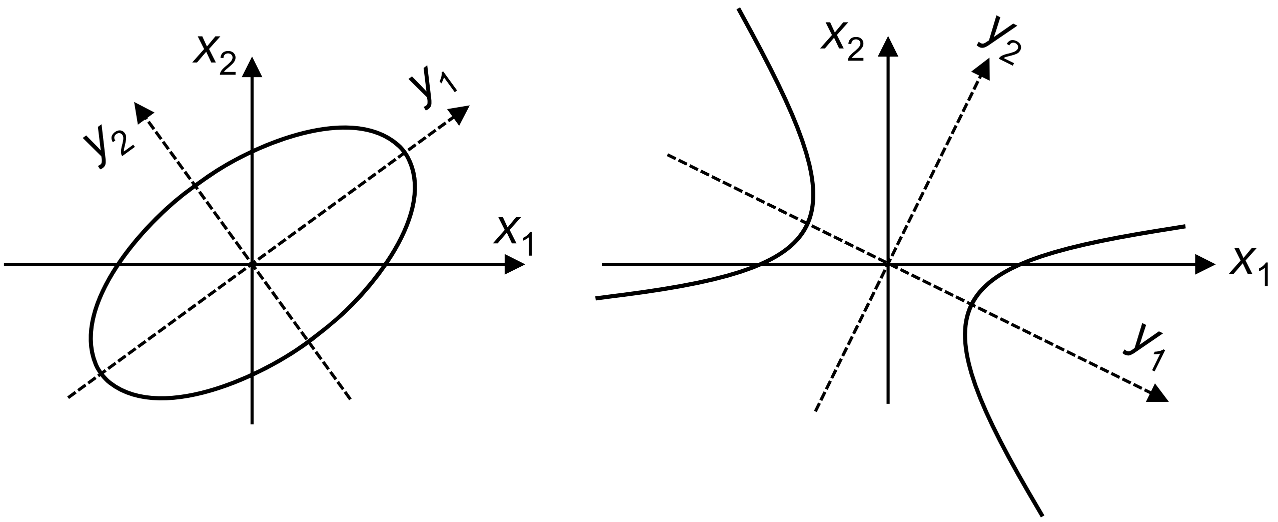
A geometric view of principal axes

- Support $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A}$ is an invertible 2×2 symmetric matrix and let c be a constant. Then any $\boldsymbol{x}$ in $\boldsymbol{R}^2$ that satisfy $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = c$ corresponds to
  - *i.* an ellipse,
  - *ii*. a circle,
  - *iii*. a hyperbola,
  - *iv*. two intersecting lines,
  - *v*. a single point, or
  - *vi*. an empty set.





- If $\boldsymbol{A}$ is a diagonal matrix, the graph of $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = c$ is in standard position.
  If $\boldsymbol{A}$ is not a diagonal matrix, the graph of $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = c$ is rotated out of standard position.



An ellipse and a hyperbola are not in standard position (out of standard position).

- Example 5.
  Find a change of variable that removes the cross-product term from $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48.$

  Solution. $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$

  Step 1. Find eigenvalues and eigenvectors of $\boldsymbol{A}$

$$\lambda_1 = 3, \quad \mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}; \quad \lambda_2 = 7, \quad \mathbf{e}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

  Step 2. Construct $\boldsymbol{P}$ and $\boldsymbol{D}$,

$$\boldsymbol{P} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \quad \boldsymbol{D} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$$

  Step 3. Transform quadratic form with no cross-product term

$$\boldsymbol{y}^T\boldsymbol{D}\boldsymbol{y} = 3y_1^2 + 7y_2^2.$$

Classifying quadratic forms

- When $\boldsymbol{A}$ is an $n \times n$ matrix, the quadratic form $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$ is a real-values function with $\boldsymbol{R}^n$.

- Definition. A quadratic form $\boldsymbol{Q}$ is:
  a. positive definite if $\boldsymbol{Q}(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq 0$,
  b. negative definite if $\boldsymbol{Q}(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \neq 0$,
  c. indefinite if $\boldsymbol{Q}(\boldsymbol{x})$ has both positive and negative values,
  d. positive semidefinite if $\boldsymbol{Q}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$,
  e. negative semidefinite if $\boldsymbol{Q}(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x}$.

- **Theorem 5. (Quadratic forms and eigenvalues)**
  Let $\boldsymbol{A}$ be an $n \times n$ symmetric matrix. Then a quadratic form $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}$ is:
  a. positive definite if and only if the eigenvalues of $\boldsymbol{A}$ are all positive.
  b. negative definite if and only if the eigenvalues of $\boldsymbol{A}$ are all negative.
  c. indefinite if and only if $\boldsymbol{A}$ has both positive and negative eigenvalues.

- Example 6. is $\boldsymbol{Q}(\boldsymbol{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite?

  Solution. $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

  The eigenvalues of $\boldsymbol{A}$ are 5, 2, and -1. So $\boldsymbol{Q}$ is an indefinite quadratic form.

- A fast way to determine whether a symmetric matrix $\boldsymbol{A}$ is positive definite.

  To factor $\boldsymbol{A}$ into the form $\boldsymbol{A} = \boldsymbol{R}^{\mathsf{T}}\boldsymbol{R}$, where $\boldsymbol{R}$ is a upper triangular with positive diagonal. Such a *Cholesky factorization* is possible if and only if $\boldsymbol{A}$ is positive definite.

- If $\boldsymbol{B}$ is $m\times n$, then $\boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}$ is positive semidefinite.
  If $\boldsymbol{B}$ is $n\times n$ and invertible, then $\boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}$ is positive definite.

  Since $\boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}$ is symmetric, $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{x} = \|\boldsymbol{B}\boldsymbol{x}\|^2 \geq 0$.

- If $n\times n$ matrix $\boldsymbol{A}$ is positive definite, then there exists a positive definite matrix $\boldsymbol{B}$ such that $\boldsymbol{A} = \boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}$.

  Since $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{\mathsf{T}}$, we can take a diagonal matrix $\boldsymbol{C}$, $\boldsymbol{D} = \boldsymbol{C}^{\mathsf{T}}\boldsymbol{C}$, and let $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{C}\boldsymbol{P}^{\mathsf{T}}$. Then

  $\boldsymbol{B}^{\mathsf{T}}\boldsymbol{B} = \boldsymbol{P}\boldsymbol{C}\boldsymbol{P}^{\mathsf{T}}\boldsymbol{P}\boldsymbol{C}\boldsymbol{P}^{\mathsf{T}} = \boldsymbol{P}\boldsymbol{C}^2\boldsymbol{P}^{\mathsf{T}} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{\mathsf{T}} = \boldsymbol{A}$.

Proof for *Cholesky factorization*

"$\Rightarrow$"

If $\boldsymbol{A} = \boldsymbol{R}^{\mathsf{T}}\boldsymbol{R}$ and $\boldsymbol{R}$ is invertible, then $\boldsymbol{A}$ is positive definite.

"$\Leftarrow$"

$\boldsymbol{A}$ is positive definite, $\boldsymbol{A} = \boldsymbol{B}^{\mathsf{T}}\boldsymbol{B}$ for some positive definite matrix $\boldsymbol{B}$.

Since the eigenvalues of $\boldsymbol{B}$ are positive and 0 is not an eigenvalue of $\boldsymbol{B}$; thus $\boldsymbol{B}$ is invertible and then the columns of $\boldsymbol{B}$ are linearly independent.

By Theorem of *QR* factorization, $\boldsymbol{B} = \boldsymbol{QR}$ for some $n\times n$ matrix $\boldsymbol{Q}$ with orthonormal columns and some upper triangular matrix $\boldsymbol{R}$ with positive entries on its diagonal.

Hence $\boldsymbol{A} = \boldsymbol{B}^{\mathsf{T}}\boldsymbol{B} = \boldsymbol{R}^{\mathsf{T}}\boldsymbol{Q}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{R}^{\mathsf{T}}\boldsymbol{R}$.

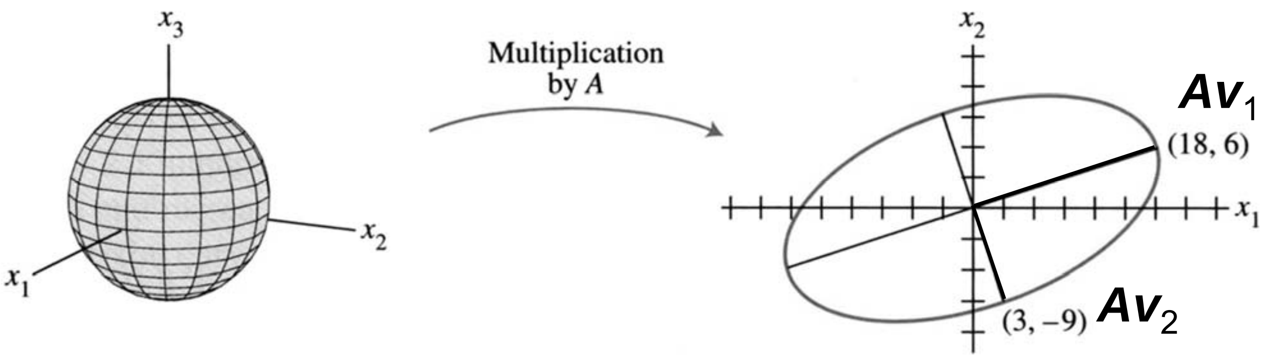
- Exercises of Section 7.2.

7.4 The singular value decomposition

- We know that not all matrices can be factored as $\boldsymbol{A} = \boldsymbol{PDP}^{-1}$ with $\boldsymbol{D}$ diagonal; however, a factorization $\boldsymbol{A} = \boldsymbol{QDP}^{-1}$ is possible for any $m \times n$ matrix $\boldsymbol{A}$. A special factorization of this type, called the singular value decomposition, is one of the most useful matrix factorizations in applied linear algebra.

- The principal of the singular value decomposition
The absolute values of the eigenvalues of a symmetric matrix $\boldsymbol{A}$ measure the amounts that $\boldsymbol{A}$ stretches or shrinks certain vectors (the eigenvectors). If $\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$ and $||\boldsymbol{x}|| = 1$, then
$$||\boldsymbol{Ax}|| = ||\lambda \boldsymbol{x}|| = |\lambda|\,||\boldsymbol{x}|| = |\lambda|. \quad (1)$$
If λ_1 is the eigenvalue with the greater magnitude, then a corresponding unit eigenvector $\boldsymbol{v}_1$ identifies a direction in which the stretching effect of $\boldsymbol{A}$ is greater.

- Example 1.
If $\boldsymbol{A} = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$, then the linear transformation $\boldsymbol{x} \mapsto \boldsymbol{Ax}$ maps the unit sphere $\{\boldsymbol{x} : ||\boldsymbol{x}|| = 1\}$ in $\boldsymbol{R}^3$ onto an ellipse in $\boldsymbol{R}^2$.



Find a unit vector $\boldsymbol{x}$ at which the length $||\boldsymbol{Ax}||$ is maximized, and compute this maximum length.
Solution.
The quantity $||\boldsymbol{Ax}||^2$ is maximized at the same $\boldsymbol{x}$ that maximizes $||\boldsymbol{Ax}||$, and $||\boldsymbol{Ax}||^2$ is easier to study.
$$||\boldsymbol{Ax}||^2 = (\boldsymbol{Ax})^T(\boldsymbol{Ax}) = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{x}.$$

A^TA is a symmetric matrix, since $(A^TA)^T = A^TA^{TT} = A^TA$. So the problem now is to maximize the quadratic form $x^T (A^TA)\, x$ subject to the constrain $||x|| = 1$.

The maximum value is the greatest eigenvalue λ_1 of A^TA. Also, the maximum value is attained at a unit eigenvector of A^TA corresponding to λ_1.

$$A^T A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}.$$

The eigenvalue s of $A^T A$ are $\lambda_1 = 360, \lambda_2 = 90$, and $\lambda_3 = 0$.

The correspond ing unit eigenvecto rs are, respective ly :

$$v_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, v_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \text{and } v_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}.$$

The maximum value of $||Ax||^2$ is 360, attained when x is the unit vector v_1. The vector Av_1 is a point on the ellipse in the above figure farthest from the origin, namely,

$$Av_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}.$$

For $\|x\| = 1$, the maximum value of $\|Ax\|$ is $\|Av_1\| = \sqrt{360} = 6\sqrt{10}$.

For the second unit eigenvector, v_2,

$$Av_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}.$$

Since v_1 and v_2 are orthogonal; $A^T A v_1 = \lambda_1 v_1$ and $A^T A v_2 = \lambda_2 v_2$.

$$Av_1 \cdot Av_2 = (Av_1)^T Av_2 = v_1^T A^T A v_2 = v_1^T (A^T A v_2) = v_1^T \lambda_2 v_2$$

$= \lambda_2 v_1^T v_2 = 0.$ Av_1 and Av_2 are also orthogonal.

#

The singular values of an $m \times n$ matrix

- Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then $\boldsymbol{A}^T\boldsymbol{A}$ is a symmetric and can be orthogonally diagonalized. Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be an orthonormal basis for $\boldsymbol{R}^n$ consisting of eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$, and let $\lambda_1, \ldots, \lambda_n$ be the associated eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$. Then for $1 \le i \le n$,

  $$\|\boldsymbol{Av}_i\|^2 = (\boldsymbol{Av}_i)^T \boldsymbol{Av}_i = \boldsymbol{v}_i^T \boldsymbol{A}^T\boldsymbol{Av}_i = \boldsymbol{v}_i^T (\lambda_i \boldsymbol{v}_i) = \lambda_i. \qquad (2)$$

  So the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ are all nonnegative. By renumbering, we can get $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n \ge 0$.

- The *singular values* of $\boldsymbol{A}$ are the square roots of the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order. That is, $\sigma_i = (\lambda_i)^{1/2}$. The singular values of $\boldsymbol{A}$ are the lengths of the vectors $\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_n$.

- Theorem 9.

  Suppose $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis of $\boldsymbol{R}^n$ consisting of eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$, arranged so that the corresponding eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ satisfy $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$, and suppose $\boldsymbol{A}$ has r nonzero singular values. Then $\{\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_r\}$ is an orthogonal basis for Col $\boldsymbol{A}$, and rank $\boldsymbol{A} = r$.

  Proof.

  Because $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ are orthogonal for $i \ne j$,

  $$(\boldsymbol{Av}_i)^T (\boldsymbol{Av}_j) = \boldsymbol{v}_i^T\boldsymbol{A}^T\boldsymbol{Av}_j = \boldsymbol{v}_i^T (\lambda_j \boldsymbol{v}_j) = 0.$$

  Thus $\{\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_n\}$ is an orthogonal set. Furthermore, since the lengths of the vectors $\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_n$ are the singular values of $\boldsymbol{A}$, and since there are r nonzero singular values, $\boldsymbol{Av}_i \ne 0$ if and only if $1 \le i \le r$. So $\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_r$ are linearly independent vectors, and they are in Col $\boldsymbol{A}$. Finally, for any $\boldsymbol{y}$ in Col $\boldsymbol{A}$, say, $\boldsymbol{y} = \boldsymbol{Ax}$, we can write $\boldsymbol{x} = c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n$, and

  $$\boldsymbol{y} = \boldsymbol{Ax} = c_1\boldsymbol{Av}_1 + \ldots + c_n\boldsymbol{Av}_n = c_1\boldsymbol{Av}_1 + \ldots + c_r\boldsymbol{Av}_r + 0 + \ldots + 0.$$

  Thus $\boldsymbol{y}$ is in *Span* $\{\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_r\}$, which shows that $\{\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_r\}$ is an orthogonal basis for Col $\boldsymbol{A}$. Hence rank $\boldsymbol{A}$ = dim Col $\boldsymbol{A} = r$.

The singular value decomposition

- The decomposition of $\boldsymbol{A}$ involves an $m\times n$ "diagonal" matrix Σ of the form

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{D} & 0 \\ 0 & 0 \end{bmatrix} \tag{3}$$

where $\boldsymbol{D}$ is an $r\times r$ diagonal matrix for some r not exceeding the smaller of m and n. (If r equals m or n or both, some or all of the zero matrices do not appear.)

- Theorem 10. (The singular value decomposition)
Let $\boldsymbol{A}$ be an $m\times n$ matrix with rank r. Then there exists an $m\times n$ matrix Σ as in Eq.(3) for which the diagonal entries in $\boldsymbol{D}$ are the first r singular values of $\boldsymbol{A}$, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$, and there exist an $m\times m$ orthogonal matrix $\boldsymbol{U}$ and an $n\times n$ orthogonal matrix $\boldsymbol{V}$ such that

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T.$$

- Any factorization $\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$, with $\boldsymbol{U}$ and $\boldsymbol{V}$ orthogonal, Σ as in Eq.(3), and positive diagonal entries in $\boldsymbol{D}$, is called a singular value decomposition (or *SVD*) of $\boldsymbol{A}$.

- The matrix $\boldsymbol{U}$ and $\boldsymbol{V}$ are not uniquely determined by $\boldsymbol{A}$, but the diagonal entries of Σ are necessarily the singular values of $\boldsymbol{A}$. The columns of $\boldsymbol{U}$ in such a decomposition are called left singular vectors of $\boldsymbol{A}$, and the columns of $\boldsymbol{V}$ are called right singular vectors of $\boldsymbol{A}$.

Proof of Theorem 10.
Let λ_i and $\boldsymbol{v}_i$ be a corresponding eigenvalue and eigenvector, so that $\{\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_r\}$ is an othogonal basis for Col $\boldsymbol{A}$. Normalize each $\boldsymbol{Av}_i$ to obtain an orthonormal basis $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$,

where $$\boldsymbol{u}_i = \frac{1}{\|\boldsymbol{Av}_i\|}\boldsymbol{Av}_i = \frac{1}{\sigma_i}\boldsymbol{Av}_i$$

and $\boldsymbol{Av}_i = \sigma_i\boldsymbol{u}_i$, $1 \leq i \leq r$. (4)

Now extend $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ to an orthonormal basis $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ of $\boldsymbol{R}^m$, and let $\boldsymbol{U} = [\boldsymbol{u}_1 \ldots \boldsymbol{u}_m]$ and $\boldsymbol{V} = [\boldsymbol{v}_1 \ldots \boldsymbol{v}_n]$.
By construction, $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices. Also from Eq.(4),

$\boldsymbol{AV} = [\boldsymbol{Av}_1 \ldots \boldsymbol{Av}_r\ \ 0\ \ldots\ 0]$ and $\boldsymbol{AV} = [\sigma_1\boldsymbol{u}_1 \ldots \sigma_r\boldsymbol{u}_r\ \ 0 \ldots\ 0]$.

Let $\boldsymbol{D}$ be the diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_r$, and let Σ be as in Eq.(3). Then

$$\begin{aligned}\boldsymbol{U\Sigma} &= [\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \cdots\ \ \boldsymbol{u}_m]\begin{bmatrix}\sigma_1 & 0 & \cdots & 0 & 0\\ 0 & \sigma_2 & \cdots & 0 & 0\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ 0 & 0 & \cdots & \sigma_r & 0\\ 0 & 0 & \cdots & 0 & 0\end{bmatrix}\\ &= [\sigma_1\boldsymbol{u}_1\ \ \cdots\ \ \sigma_r\boldsymbol{u}_r\ \ 0\ \ \cdots\ \ 0]\\ &= \boldsymbol{AV}.\end{aligned}$$

Since $\boldsymbol{V}$ is an orthogonal matrix, $\boldsymbol{U\Sigma V}^T = \boldsymbol{AVV}^T = \boldsymbol{A}$.

❁ Example 3.

Using the results of Example 1 to construct a singular value decomposition of $\boldsymbol{A} = \begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$.

Solution.

A construction can be divided into three steps.

Step 1. Find an orthogonal diagonalization of $\boldsymbol{A}^T\boldsymbol{A}$. That is to find the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ and a corresponding orthonormal set of eigenvectors.

Step 2. Set up $\boldsymbol{V}$ and Σ.

$$\boldsymbol{V} = [\boldsymbol{v}_1\ \ \boldsymbol{v}_2\ \ \boldsymbol{v}_3] = \begin{bmatrix}1/3 & -2/3 & 2/3\\ 2/3 & -1/3 & -2/3\\ 2/3 & 2/3 & 1/3\end{bmatrix}.$$

$$\boldsymbol{D} = \begin{bmatrix}6\sqrt{10} & 0\\ 0 & 3\sqrt{10}\end{bmatrix},\ \boldsymbol{\Sigma} = [\boldsymbol{D}\ \ 0] = \begin{bmatrix}6\sqrt{10} & 0 & 0\\ 0 & 3\sqrt{10} & 0\end{bmatrix}.$$

Step 3. Construct $\boldsymbol{U}$.

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1 = \frac{1}{6\sqrt{10}}\begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$\boldsymbol{u}_2 = \frac{1}{\sigma_2}\boldsymbol{A}\boldsymbol{v}_2 = \frac{1}{3\sqrt{10}}\begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

and $\boldsymbol{U} = [\boldsymbol{u}_1 \ \ \boldsymbol{u}_2]$.

Thus the singular value decomposition of $\boldsymbol{A}$ is

$$\boldsymbol{A} = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}.$$

- Example 4.

Find a singular value decomposition of $\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Solution.

Step 1. Find the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ and a corresponding orthonormal set of eigenvectors.

$$\lambda_1 = 18,\ \boldsymbol{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \text{ and } \lambda_2 = 0,\ \boldsymbol{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Step 2. Set up $\boldsymbol{V}$ and Σ.

$$\boldsymbol{V} = [\boldsymbol{v}_1 \ \ \boldsymbol{v}_2] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}. \qquad \boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}.$$

$$\boldsymbol{D} = \sqrt{18} = 3\sqrt{2}.$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{D} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Step 3. Construct $\boldsymbol{U}$.

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1 = \frac{1}{3\sqrt{2}}\begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}, \text{ but } \boldsymbol{A}\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

thus we need two orthogonal unit vectors $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ that are orthogonal to $\boldsymbol{u}_1$. That is, we want to find two vectors $\boldsymbol{x}$ such that $\boldsymbol{u}_1^T\boldsymbol{x} = 0$. The problem is equivalent to the equation $x_1 - 2x_2 + 2x_3 = 0$.
A basis for the solution set is

$$\boldsymbol{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{w}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Apply the Gram – Schmidt process to $\{\boldsymbol{w}_2, \boldsymbol{w}_3\}$ to get

$$\boldsymbol{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} \text{ and } \boldsymbol{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}.$$

Finally, set $\boldsymbol{U} = [\boldsymbol{u}_1\ \boldsymbol{u}_2\ \boldsymbol{u}_3]$.

$$\text{Then } \boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix}\begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

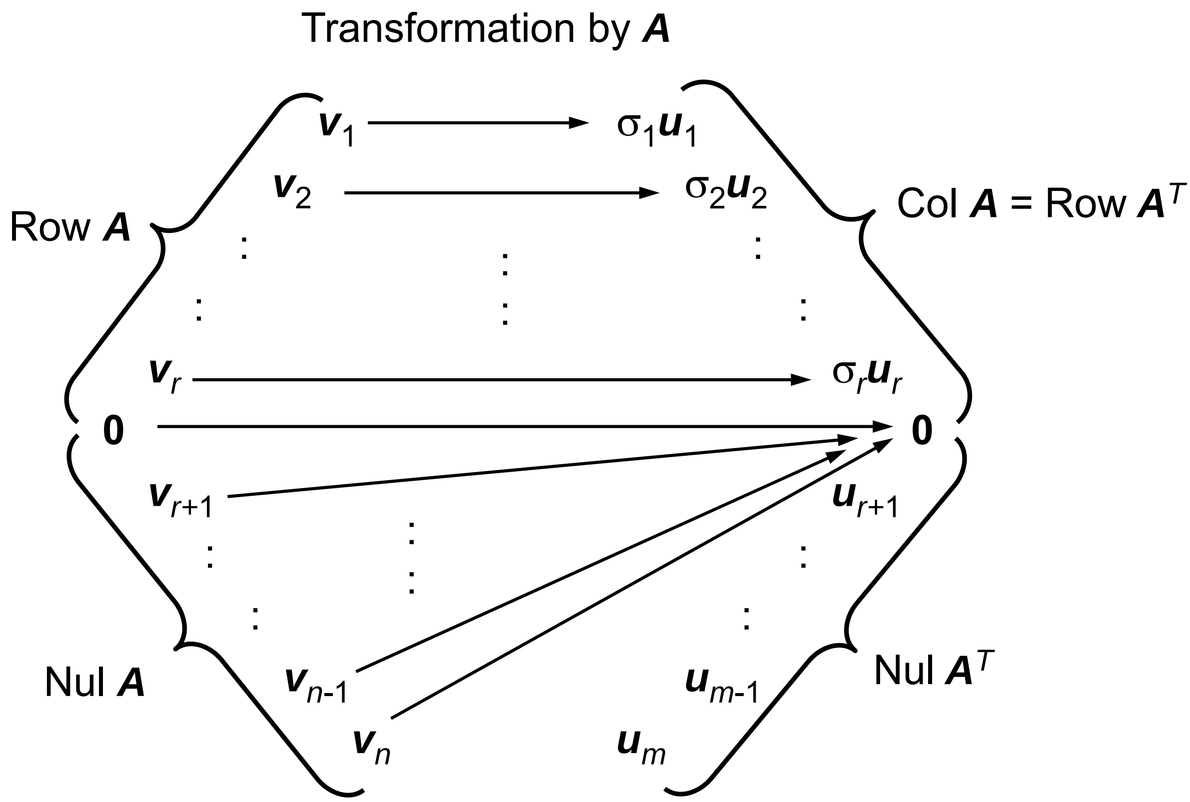
Applications of the singular value decomposition

- The *SVD* is often used to estimate the rank of a matrix, as note above. Several other numerical applications are described briefly below, and an application to image processing is presented in my course of Image Processing.
- Example 5. (The condition number)
  Consider the equation $\boldsymbol{Ax} = \boldsymbol{b}$. If the singular values of $\boldsymbol{A}$ are extremely large or small, roundoff errors are almost inevitable, but an error analysis is aided by knowing the entries in Σ and $\boldsymbol{V}$. If $\boldsymbol{A}$ is an invertible $n\times n$ matrix, then the ratio σ_1/σ_n of the largest and smallest singular values given the *condition number* of $\boldsymbol{A}$.

- Example 6. (*Bases for fundamental subspaces*)
  Given an *SVD* for an $m \times n$ matrix $\boldsymbol{A}$, let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ be the left singular vectors, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be the right singular vectors, $\sigma_1, \ldots, \sigma_n$ be the singular values, and r be the rank of $\boldsymbol{A}$.

  i. By Theorem 9, $\{\boldsymbol{u}_1, ., \boldsymbol{u}_r\}$ is an orthonormal basis for Col $\boldsymbol{A}$.
  ii. Since (Col $\boldsymbol{A})^{\perp}$ = Nul $\boldsymbol{A}^T$, $\{\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m\}$ is an orthonormal basis for Nul $\boldsymbol{A}^T$.
  iii. Since $\|\boldsymbol{A}\boldsymbol{v}_i\| = \sigma_i$ for $1 \le i \le n$, and σ_i is 0 if and only if $i > r$, the vectors $\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$ span a subspace of Nul $\boldsymbol{A}$ of dimension $n - r$. By the Rank Theorem, dim Nul $\boldsymbol{A} = n -$ rank $\boldsymbol{A}$. It follows that $\{\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis for Nul $\boldsymbol{A}$.
  iv. Since (Nul $\boldsymbol{A}^T)^{\perp}$ = Col $\boldsymbol{A}$, interchanging $\boldsymbol{A}$ and $\boldsymbol{A}^T$, we have (Nul $\boldsymbol{A})^{\perp}$ = Col $\boldsymbol{A}^T$ = Row $\boldsymbol{A}$; thus $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ is an orthonormal basis for Row $\boldsymbol{A}$.

- The four fundamental subspaces and the action of $\boldsymbol{A}$ are diagrammed as follows.



- Theorem. (The invertible matrix theorem (concluded)
  Let $\boldsymbol{A}$ be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that $\boldsymbol{A}$ is an invertible matrix.
  u. $(\text{Col } \boldsymbol{A})^{\perp} = \{0\}$.
  v. $(\text{Nul } \boldsymbol{A})^{\perp} = \boldsymbol{R}^{n}$.
  w. $\text{Row } \boldsymbol{A} = \boldsymbol{R}^{n}$.
  x. $\boldsymbol{A}$ has n nonzero singular values.

- Example 7. (*Reduced SVD and the pseudoinverse of **A***)
  When Σ contains rows or columns of zeros, a more compact decomposition of $\boldsymbol{A}$ is possible. Using the notation established above, let r = rank $\boldsymbol{A}$, and partition $\boldsymbol{U}$ and $\boldsymbol{V}$ into submatrices whose first blocks contain r columns:
  $\boldsymbol{U} = [\, \boldsymbol{U}_r \;\; \boldsymbol{U}_{m-r} \,]$, where $\boldsymbol{U}_r = [\, \boldsymbol{u}_1 \dots \boldsymbol{u}_r \,]$ and
  $\boldsymbol{V} = [\, \boldsymbol{V}_r \;\; \boldsymbol{V}_{n-r} \,]$, where $\boldsymbol{V}r = [\, \boldsymbol{v}_1 \dots \boldsymbol{v}_r \,]$ and
  Then $\boldsymbol{U}_r$ is $m \times r$ and $\boldsymbol{V}_r$ is $n \times r$. Then partitioned matrix multiplication shows that

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{U}_r & \boldsymbol{U}_{m-r} \end{bmatrix} \begin{bmatrix} \boldsymbol{D} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_r^T \\ \boldsymbol{V}_{n-r}^T \end{bmatrix} = \boldsymbol{U}_r \boldsymbol{D} \boldsymbol{V}_r^T . \qquad (9)$$

  This factorization of $\boldsymbol{A}$ is called a reduced singular value decomposition of $\boldsymbol{A}$. Since the diagonal entries in $\boldsymbol{D}$ are nonzero, we can form the following matrix, called the *pseudoinverse* (also, *Moore-Psnrose inverse*) of $\boldsymbol{A}$:

$$\boldsymbol{A}^{+} = \boldsymbol{V}_r \boldsymbol{D}^{-1} \boldsymbol{U}_r^{T}. \qquad (10)$$

- Example 8. (*Least-square solution*)
  Given the equation $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$, use the pseudo inverse of $\boldsymbol{A}$ to define $\hat{\boldsymbol{x}} = \boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{V}_r \boldsymbol{D}^{-1} \boldsymbol{U}_r^T \boldsymbol{b}$.
  Then, from the *SVD* in Eq.(9),

$$\begin{aligned} \boldsymbol{A}\hat{\boldsymbol{x}} &= (\boldsymbol{U}_r \boldsymbol{D}^{-1} \boldsymbol{V}_r^T)(\boldsymbol{V}_r \boldsymbol{D}^{-1} \boldsymbol{U}_r^T \boldsymbol{b}) \\ &= \boldsymbol{U}_r \boldsymbol{D} \boldsymbol{D}^{-1} \boldsymbol{U}_r^T \boldsymbol{b} \qquad (\because\ \boldsymbol{V}_r^T \boldsymbol{V}_r = I_r) \\ &= \boldsymbol{U}_r \boldsymbol{U}_r^T \boldsymbol{b}. \end{aligned}$$

  Where $\boldsymbol{U}_r \boldsymbol{U}_r^T \boldsymbol{b}$ is the orthogonal projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ onto Col $\boldsymbol{A}$.
  Thus $\hat{\boldsymbol{x}}$ is a least – square solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. In fact, this $\hat{\boldsymbol{x}}$ has the smallest length among all least – square solution $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

- Exercises of Section 7.4.

- A $n \times n$ square matrix exactly has n eigenvalues,
  but may has less than or equal to n eigenvectors.
- For a general square matrix,
  eigenvalues can be real or complex number, and
  eigenvectors may be linearly dependent or independent.
- If a matrix has eigenvalue 0, the matrix is not invertible.
  If a $n \times n$ matrix has n *LI* eigenvector, the matrix is diagonalizable.
- For a symmetric matrix, the eigenvalues must be real number, the eigenvectors must be orthogonalized.
- For $\boldsymbol{A}^T\boldsymbol{A}$ matrix, the eigenvalues must be nonnegative,
  the eigenvectors must be orthogonalized.