

# Chapter 8. Symmetric matrices and quadratic forms

C.O.S. Sorzano

Biomedical Engineering

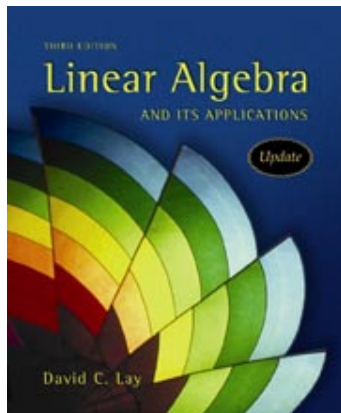
December 3, 2013



CEU

*Universidad  
San Pablo*

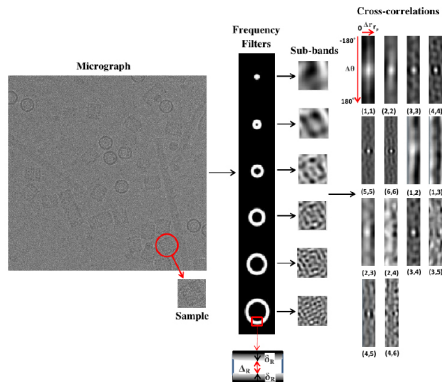
- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - Constrained optimization (b)
  - Singular Value Decomposition (SVD) (c)



D. Lay. Linear algebra and its applications (3rd ed). Pearson (2006). Chapter 7.

# Applications

In this example of particle picking in Single Particles, one of the features we analyze is the autocorrelation function at different subbands. The autocorrelation is a symmetric matrix.



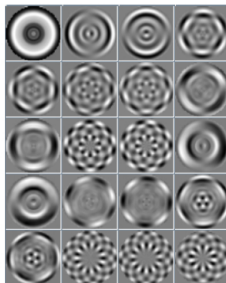
V. Abrishami, A. Zaldívar-Peraza, J.M. de la Rosa-Trevín, J. Vargas, J. Otón, R. Marabini, Y. Shkolnisky, J.M. Carazo, C.O.S. Sorzano. *A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs*. *Bioinformatics* (2013)

# Applications

In one of the steps, we construct a basis that spans the set of rotations of the particle template. For doing so, perform a Principal Component Analysis that diagonalizes the covariance matrix (which is again a symmetric matrix).



(a)



V. Abrishami, A. Zaldívar-Peraza, J.M. de la Rosa-Trevín, J. Vargas, J. Otón, R. Marabini, Y. Shkolnisky, J.M. Carazo, C.O.S. Sorzano. *A pattern matching approach to the automatic selection of particles from low-contrast electron micrographs*. *Bioinformatics* (2013)

- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - Constrained optimization (b)
  - Singular Value Decomposition (SVD) (c)

# Diagonalization of symmetric matrices

## Definition 1.1 (Symmetric matrix)

$A \in \mathcal{M}_{n \times n}$  is a **symmetric matrix** iff  $A = A^T$ .

## Example

The following two matrices are symmetric

$$\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{pmatrix}$$

## Example

Let's diagonalize the matrix  $A = \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}$  The characteristic equation is

$$|A - \lambda I| = 0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

# Diagonalization of symmetric matrices

The associated eigenvectors are

$$\begin{aligned}\lambda = 8 \quad \mathbf{v}_1 &= (-1, 1, 0) \rightarrow \mathbf{u}_1 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) \\ \lambda = 6 \quad \mathbf{v}_2 &= (-1, -1, 2) \rightarrow \mathbf{u}_2 = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) \\ \lambda = 3 \quad \mathbf{v}_3 &= (1, 1, 1) \rightarrow \mathbf{u}_3 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\end{aligned}$$

The  $\mathbf{v}$  vectors constitute an orthogonal basis of  $\mathbb{R}^3$  and after normalizing them ( $\mathbf{u}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ ), we have an orthonormal basis. Thus, we can factorize  $A$  as  $A = PDP^{-1}$  with

$$P = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \quad D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Exploiting the fact that  $P$  is orthonormal, then  $P^{-1} = P^T$  and  $A = PDP^T$ .



# Diagonalization of symmetric matrices

## Theorem 1.1

If  $A$  is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

Proof

Let  $\mathbf{v}_1$  and  $\mathbf{v}_2$  be two eigenvectors from two different eigenvalues  $\lambda_1$  and  $\lambda_2$ . Let's show that  $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$

$$\begin{aligned}\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 && \text{[By definition]} \\ &= (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{[Definition of eigenvector]} \\ &= \mathbf{v}_1^T A^T \mathbf{v}_2 && \text{[Transpose of product]} \\ &= \mathbf{v}_1^T (A\mathbf{v}_2) && \text{[A is symmetric]} \\ &= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{[Definition of eigenvector]} \\ &= \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2) && \text{[By definition]}\end{aligned}$$

Hence  $(\lambda_1 - \lambda_2)(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0$  but  $\lambda_1 - \lambda_2 \neq 0$  because the two eigenvalues are different. Consequently,  $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$  (q.e.d.)

# Diagonalization of symmetric matrices

## Definition 1.2 (Orthogonal diagonalization)

$A$  is **orthogonally diagonalizable** iff  $A = PDP^T$  being  $P$  an orthogonal (i.e.,  $P^{-1} = P^T$ ).

## Theorem 1.2

$A$  is orthogonally diagonalizable iff  $A$  is symmetric.

Proof orthogonally diagonalizable  $\Rightarrow$  symmetric

Let us assume that  $A = PDP^T$ , then

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PD^T P^T = PDP^T = A$$

Proof orthogonally diagonalizable  $\Leftarrow$  symmetric

We omit this proof since it is more difficult.

# Diagonalization of symmetric matrices

## Example

Let's orthogonally diagonalize  $A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ .

### Solution

The characteristic equation is

$$|A - \lambda I| = 0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

Its associated eigenvectors are

$$\begin{aligned} \lambda = 7 \quad \mathbf{v}_1 = (1, 0, 1) &\rightarrow \mathbf{u}_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \\ &\quad \mathbf{v}_2 = \left(-\frac{1}{2}, 1, 2\right) \rightarrow \mathbf{u}_2 = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right) \\ \lambda = -2 \quad \mathbf{v}_3 = \left(-1, -\frac{1}{2}, 1\right) &\rightarrow \mathbf{u}_3 = \left(-\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right) \end{aligned}$$

$\mathbf{u}_1$  and  $\mathbf{u}_2$  are unitary and linearly independent, but they are not orthogonal.  $\mathbf{u}_3$  is orthogonal to the other two vectors because it belongs to a different eigenspace (see Theorem 1.1).

# Diagonalization of symmetric matrices

We can orthogonalize  $\mathbf{u}_1$  and  $\mathbf{u}_2$  following the Gram-Schmidt procedure:

$$\mathbf{w}_1 = \mathbf{v}_1 = \left( \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)$$

$$\mathbf{w}'_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1 = \left( -\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0 \right) - \left( -\frac{1}{\sqrt{10}} \right) \left( \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) = \left( -\frac{1}{2\sqrt{5}}, \frac{2}{\sqrt{5}}, \frac{1}{2\sqrt{5}} \right)$$

$$\mathbf{w}_2 = \frac{\mathbf{w}'_2}{\|\mathbf{w}'_2\|} = \left( -\frac{1}{3\sqrt{2}}, \frac{2\sqrt{2}}{3}, \frac{1}{3\sqrt{2}} \right)$$

$$\mathbf{w}_3 = \mathbf{v}_3 = \left( -\frac{2}{3}, -\frac{1}{3}, \frac{2}{3} \right)$$

So  $A = PDP^T$  with

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{2\sqrt{2}}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \end{pmatrix} \quad D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

# Diagonalization of symmetric matrices

## Definition 1.3 (Spectrum of a matrix)

*The set of eigenvalues of a matrix is called the **spectrum** of that matrix.*

## Theorem 1.3 (Spectral theorem for symmetric matrices)

*An  $n \times n$  symmetric matrix has the following properties:*

- 1 *A has  $n$  real eigenvalues (including multiplicities).*
- 2 *The dimension of each eigenspace is the multiplicity of the corresponding eigenvalue as root of the characteristic equation.*
- 3 *Eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.*
- 4 *A is orthogonally diagonalizable.*

# Diagonalization of symmetric matrices

## Definition 1.4 (Spectral decomposition of symmetric matrices)

Let  $A = PDP^T$  with  $P = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n)$ . Then

$$\begin{aligned} A &= (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_n) \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \dots \\ \mathbf{u}_n^T \end{pmatrix} \\ &= (\lambda_1 \mathbf{u}_1 \ \lambda_2 \mathbf{u}_2 \ \dots \ \lambda_n \mathbf{u}_n) \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \dots \\ \mathbf{u}_n^T \end{pmatrix} \\ &= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T \end{aligned}$$

The latest equation is the **spectral decomposition** of  $A$ . Each one of the terms  $\lambda_i \mathbf{u}_i \mathbf{u}_i^T$  is an  $n \times n$  matrix of rank 1 (since all the columns are multiples of  $\mathbf{u}_i$ ). Additionally,  $\mathbf{u}_i \mathbf{u}_i^T \mathbf{x}$  is the orthogonal projection of any vector onto the subspace generated by  $\mathbf{u}_i$ .

# Diagonalization of symmetric matrices

## Example

Write the spectral decomposition of

$$A = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

### Solution

Consider  $\mathbf{u}_1 = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$  be the first column of  $P$  and  $\mathbf{u}_2 = (-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ . Then

$$\mathbf{u}_1 \mathbf{u}_1^T = \begin{pmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix} \quad \mathbf{u}_2 \mathbf{u}_2^T = \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{pmatrix}$$

The spectral decomposition is therefore

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T = 8 \begin{pmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix} + 3 \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{pmatrix}$$

## Exercises

From Lay (3rd ed.), Chapter 7, Section 1:

- 7.1.6
- 7.1.7
- 7.1.13
- 7.1.23
- 7.1.27
- 7.1.29
- 7.1.35



- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - Constrained optimization (b)
  - Singular Value Decomposition (SVD) (c)

# Quadratic forms

## Introduction

Most expressions appearing so far are linear:  $A\mathbf{x}$ ,  $\langle \mathbf{w}, \mathbf{x} \rangle$ ,  $\mathbf{x}^T$ , that is, if we construct an operator  $T(\mathbf{x})$  with them (e.g.,  $T(\mathbf{x}) = A\mathbf{x}$ ,  $T(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle$ ,  $T(\mathbf{x}) = \mathbf{x}^T$ ), it meets

$$T(a\mathbf{x}_1 + b\mathbf{x}_2) = aT(\mathbf{x}_1) + bT(\mathbf{x}_2)$$

However, there are nonlinear expressions like  $\mathbf{x}^T \mathbf{x}$ . Particularly, this one is said to be quadratic and they normally appear in applications of linear algebra to engineering (like optimization) and signal processing (like signal power). They also arise in physics (as potential and kinetic energy), differential geometry (as the normal curvature of surfaces) and statistics (as confidence ellipsoids).

# Quadratic forms

## Definition 2.1 (Quadratic forms)

A quadratic form in  $\mathbb{R}^n$  is a function  $Q(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$  that can be computed as

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

being  $A \in \mathcal{M}_{n \times n}$  a symmetric matrix.

## Example

$$\textcircled{1} \quad Q(\mathbf{x}) = \mathbf{x}^T I \mathbf{x} = (x_1 \quad x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$$

$$\textcircled{2} \quad Q(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} \mathbf{x} = 4x_1^2 + 3x_2^2$$

$$\textcircled{3} \quad Q(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix} \mathbf{x} = 3x_1^2 + 7x_2^2 - 4x_1x_2$$

$$\textcircled{4} \quad Q(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 5 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 3 & 4 \\ 0 & 4 & 2 \end{pmatrix} \mathbf{x} = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$$

# Change of variables in quadratic forms

## Change of variables

A change of variables is an equation of the form  $\mathbf{x} = P\mathbf{y}$  or equivalently  $P^{-1}\mathbf{x} = \mathbf{y}$ , where  $P$  is an invertible matrix. Exploiting the fact that, in a quadratic form,  $A$  is symmetric, then we have  $A = PDP^T$ . We perform the change of variables

$$\mathbf{x} = P\mathbf{y}$$

to obtain

$$Q(\mathbf{x}) = (P\mathbf{y})^T A (P\mathbf{y}) = \mathbf{y}^T P^T A P \mathbf{y} = Q(\mathbf{y})$$

But we know

$$A = PDP^T \Rightarrow D = P^T A P$$

Consequently

$$Q(\mathbf{y}) = \mathbf{y}^T D \mathbf{y}$$

That is, there is a basis, in which the matrix of the quadratic form is diagonal.

# Change of variables in quadratic forms

## Example

Consider  $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$  with

$$A = \begin{pmatrix} 1 & -4 \\ -4 & -5 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -7 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

That is

$$Q(\mathbf{x}) = x_1^2 - 5x_2^2 - 8x_1x_2$$

If we make the change of variable

$$\mathbf{y} = P^T \mathbf{x} = \begin{pmatrix} \frac{2}{\sqrt{5}}x_1 - \frac{1}{\sqrt{5}}x_2 \\ \frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_2 \end{pmatrix}$$

then

$$Q(\mathbf{y}) = \mathbf{y}^T D \mathbf{y} = 3y_1^2 - 7y_2^2$$

## Change of variables in quadratic forms

Let's check that effectively both ways of calculating the quadratic form are equivalent. For doing so, we'll calculate the value of  $Q(\mathbf{x})$  for  $\mathbf{x} = (2, -2)$ :

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = 2^2 - 5 \cdot (-2)^2 - 8 \cdot 2 \cdot (-2) = 4 - 20 + 32 = 16$$

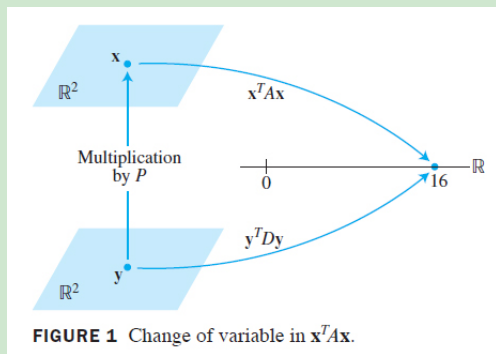
If we make the change of variable

$$\mathbf{y} = \begin{pmatrix} \frac{2}{\sqrt{5}} 2 - \frac{1}{\sqrt{5}} (-2) \\ \frac{1}{\sqrt{5}} 2 + \frac{2}{\sqrt{5}} (-2) \end{pmatrix} = \begin{pmatrix} \frac{6}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$$

then

$$Q(\mathbf{y}) = \mathbf{y}^T D \mathbf{y} = 3 \left( \frac{6}{\sqrt{5}} \right)^2 - 7 \left( -\frac{2}{\sqrt{5}} \right)^2 = 3 \frac{36}{5} - 7 \frac{4}{5} = \frac{80}{5} = 16$$

# Change of variables in quadratic forms



## Theorem 2.1 (Principal axes theorem)

Let  $A \in \mathcal{M}_{n \times n}$  be a symmetric matrix. Then, there exists a change of variable  $\mathbf{x} = P\mathbf{y}$  such that the quadratic form  $\mathbf{x}^T A \mathbf{x}$  becomes  $\mathbf{y}^T D \mathbf{y}$  with  $D$  an  $n \times n$  diagonal matrix. The columns of  $P$  are the principal axes.

# Principal axes

## A geometric view of the principal axes

Consider the quadratic form  $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$  with  $\mathbf{x} \in \mathbb{R}^2$  and the isocurve  $Q(\mathbf{x}) = c$ . The isocurve is either an ellipse, a circle, a hyperbola, two intersecting lines, a point, or contains no points at all. If  $A$  is diagonal, then

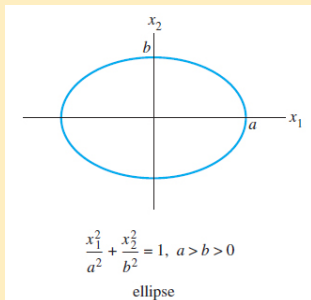
$$Q(\mathbf{x}) = a_{11}x_1^2 + a_{22}x_2^2 = c$$

The equation of an ellipse is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

with  $a, b > 0$ . Therefore

$$a = \sqrt{\frac{c}{a_{11}}} \quad b = \sqrt{\frac{c}{a_{22}}}$$



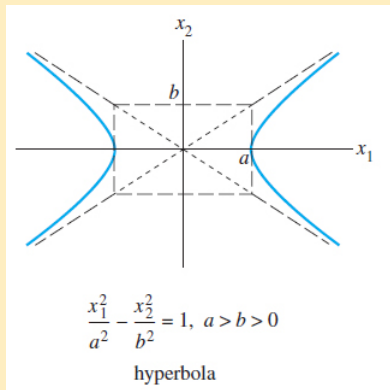


# Principal axes

The equation of a hyperbola is

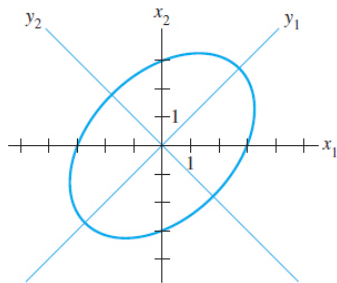
$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1$$

with  $a, b > 0$

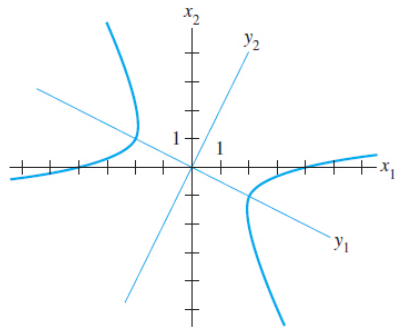


# Principal axes

If  $A$  is not diagonal, then the ellipse or the hyperbola are rotated



(a)  $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$



(b)  $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

# Principal axes

## Example

Let's analyze the rotated ellipse

$$5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$$

The corresponding matrix is

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

So,

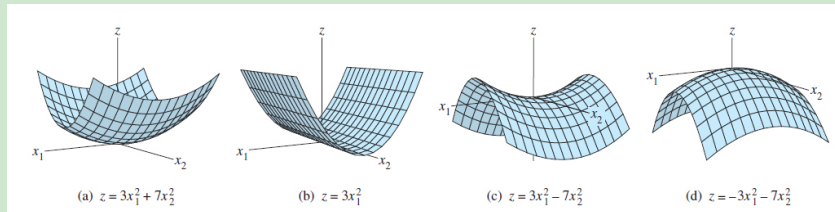
$$a = \sqrt{\frac{c}{a_{11}}} = \sqrt{\frac{48}{3}} = 3 \quad b = \sqrt{\frac{c}{a_{22}}} = \sqrt{\frac{48}{7}} \approx 2.65$$

The change of variable  $\mathbf{x} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{y}$  diagonalizes the quadratic form (see the new axes in the previous slide).

# Classification of quadratic forms

## Example

Look at the following surfaces defined as  $z = Q(\mathbf{x})$



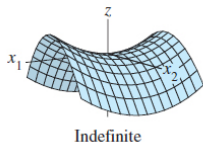
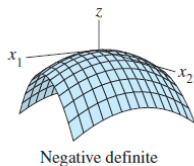
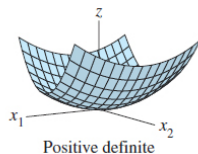
The curves seen in  $\mathbb{R}^2$  are the cut of these surfaces with the plane  $z = c$ . It is obvious that some of the surfaces are always above  $z = 0$  (a and b), others are always below  $z = 0$  (d), and still other are sometimes below and sometimes above  $z = 0$  (c).

# Classification of quadratic forms

## Definition 2.2 (Classification of quadratic forms)

We say  $Q(\mathbf{x})$  is

- **positive definite** if
$$Q(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$$
- **negative definite** if
$$Q(\mathbf{x}) < 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$$
- **indefinite** if  $Q(\mathbf{x})$  assumes both positive and negative values
- **positive semidefinite** if
$$Q(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$$
- **negative semidefinite** if
$$Q(\mathbf{x}) \leq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$$



# Classification of quadratic forms

## Theorem 2.2 (Classification of quadratic forms and quadratic forms)

Let  $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$  with  $A \in \mathcal{M}_{n \times n}$  and symmetric. Let  $\lambda_i$  be the eigenvalues of  $A$ .  $Q(\mathbf{x})$  is

- **positive definite** iff  $\lambda_i > 0 \quad \forall i$
- **negative definite** iff  $\lambda_i < 0 \quad \forall i$
- **indefinite** iff there are positive and negative eigenvalues
- **positive semidefinite** iff  $\lambda_i \geq 0 \quad \forall i$
- **negative semidefinite** iff  $\lambda_i \leq 0 \quad \forall i$

### Proof

By the Theorem of Principal Axes (Theorem 2.1), there is a change of variable such that

$$Q(\mathbf{y}) = \mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

where  $\lambda_i$  is the  $i$ -th eigenvalue. The values of  $Q$  depend on  $\lambda_i$  in the way that the theorem states (e.g.,  $\forall \mathbf{y} \neq \mathbf{0} \quad Q(\mathbf{y}) > 0$  iff  $\lambda_i > 0 \quad \forall i$ , etc.)

# Classification of quadratic forms

## Examples

- $Q(\mathbf{x}) = 3x_1^2 + 7x_2^2$  is positive definite because its eigenvalues are 3 and 7 (both larger than 0).
- $Q(\mathbf{x}) = 3x_1^2$  is positive semidefinite because its eigenvalues are 3 and 0 (both larger or equal than 0).
- $Q(\mathbf{x}) = 3x_1^2 - 7x_2^2$  is indefinite because its eigenvalues are 3 and -7 (one positive and another negative).
- $Q(\mathbf{x}) = -3x_1^2 - 7x_2^2$  is negative definite because its eigenvalues are -3 and -7 (both smaller than 0).

## Definition 2.3 (Classification of symmetric matrices)

A symmetric **matrix is positive definite** if its corresponding quadratic form is positive definite. Analogously for the rest of the classification.

# Classification of quadratic forms

## Cholesky factorization

Cholesky factorization factorizes a symmetric matrix  $A$  as

$$A = R^T R$$

being  $R$  an upper triangular matrix.  $A$  is positive definite if all entries in the diagonal of  $R$  are positive.



## Exercises

From Lay (3rd ed.), Chapter 7, Section 2:

- 7.2.1
- 7.2.3
- 7.2.5
- 7.2.7
- 7.2.19
- 7.2.23
- 7.2.24
- 7.2.26
- 7.2.27

- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - **Constrained optimization (b)**
  - Singular Value Decomposition (SVD) (c)

# Constrained optimization

## Introduction

Many problems in engineering or physics are of the form

$$\begin{array}{ll} \min & Q(\mathbf{x}) \\ \text{subject to} & \|\mathbf{x}\|^2 = 1 \end{array} \quad \text{or} \quad \begin{array}{ll} \max & Q(\mathbf{x}) \\ \text{subject to} & \|\mathbf{x}\|^2 = 1 \end{array}$$

## Example

Calculate the minimum and maximum of  $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$  subject to  $\|\mathbf{x}\|^2 = 1$ . Solution

By taking the minimum and maximum coefficient in  $Q(\mathbf{x})$  we have

$$\begin{aligned} 3x_1^2 + 3x_2^2 + 3x_3^2 &\leq Q(\mathbf{x}) \leq 9x_1^2 + 9x_2^2 + 9x_3^2 \\ 3(x_1^2 + x_2^2 + x_3^2) &\leq Q(\mathbf{x}) \leq 9(x_1^2 + x_2^2 + x_3^2) \\ 3 &\leq Q(\mathbf{x}) \leq 9 \end{aligned}$$

The minimum value  $Q(\mathbf{x}) = 3$  is attained for  $\mathbf{x} = (0, 0, 1)$ , while the maximum value  $Q(\mathbf{x}) = 9$  is attained for  $\mathbf{x} = (1, 0, 0)$ . In fact the minimum and maximum values that the constrained quadratic form can take are  $\lambda_{\min}$  and  $\lambda_{\max}$ .

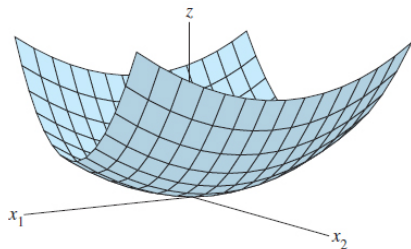
# Constrained optimization

## Example

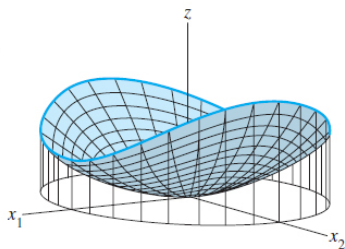
Calculate the minimum and maximum of  $Q(\mathbf{x}) = 3x_1^2 + 7x_2^2$  subject to  $\|\mathbf{x}\|^2 = 1$ .

### Solution

$\|\mathbf{x}\|^2 = 1$  is a cylinder in  $\mathbb{R}^3$  while  $z = Q(\mathbf{x})$  is a parabolic surface. The minimum and maximum of the constrained problem are attained among those points belonging to the curve that is the intersection of both surfaces.



**FIGURE 1**  $z = 3x_1^2 + 7x_2^2$ .



**FIGURE 2** The intersection of  $z = 3x_1^2 + 7x_2^2$  and the cylinder  $x_1^2 + x_2^2 = 1$ .

# Constrained optimization

## Theorem 3.1

Let  $A$  be a symmetric matrix and let

$$m = \min \{ \mathbf{x}^T A \mathbf{x} \mid \|\mathbf{x}\|^2 = 1 \}$$
$$M = \max \{ \mathbf{x}^T A \mathbf{x} \mid \|\mathbf{x}\|^2 = 1 \}$$

Then,  $M = \lambda_{\max}$  and  $m = \lambda_{\min}$ .  $M$  is attained for  $\mathbf{x} = \mathbf{u}_{\max}$  (the eigenvector associated to  $\lambda_{\max}$ ) and  $m$  is attained for  $\mathbf{x} = \mathbf{u}_{\min}$  (the eigenvector associated to  $\lambda_{\min}$ ).

Proof

Let's orthogonally diagonalize  $A$  as  $A = PDP^T$  and we make the change variables  $\mathbf{y} = P^T \mathbf{x}$ . We already know that

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \mathbf{y}^T D \mathbf{y}$$

Additionally  $\|\mathbf{y}\|^2 = \|\mathbf{x}\|^2$  because

$$\|\mathbf{y}\|^2 = \mathbf{y}^T \mathbf{y} = (P^T \mathbf{x})^T (P^T \mathbf{x}) = \mathbf{x}^T P P^T \mathbf{x} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2$$

In particular  $\|\mathbf{y}\| = 1 \Leftrightarrow \|\mathbf{x}\| = 1$ .

# Constrained optimization

Then,

$$m = \min \{ \mathbf{y}^T D \mathbf{y} \mid \|\mathbf{y}\|^2 = 1 \}$$
$$M = \max \{ \mathbf{y}^T D \mathbf{y} \mid \|\mathbf{y}\|^2 = 1 \}$$

Since  $D$  is diagonal we have

$$\mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

Let's look for the maximum of these values subject to  $\|\mathbf{y}\| = 1$ . Consider the maximum eigenvalue,  $\lambda_{\max}$ , then

$$\begin{aligned} \mathbf{y}^T D \mathbf{y} &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 \\ &\leq \lambda_{\max} y_1^2 + \lambda_{\max} y_2^2 + \dots + \lambda_{\max} y_n^2 \\ &= \lambda_{\max} (y_1^2 + y_2^2 + \dots + y_n^2) \\ &= \lambda_{\max} \|\mathbf{y}\|^2 = \lambda_{\max} \end{aligned}$$

# Constrained optimization

In fact the value  $\lambda_{max}$  is attained for  $\mathbf{y}_{max} = (0 \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0)$ , where the 1 is at the location corresponding to  $\lambda_{max}$ . The corresponding  $\mathbf{x}$  is

$$\mathbf{x} = P\mathbf{y} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_{max-1} \ \mathbf{u}_{max} \ \mathbf{u}_{max+1} \ \dots \ \mathbf{u}_n) \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \mathbf{u}_{max}$$

We could reason analogously for the minimum.

# Constrained optimization

## Example

Let  $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}$ . Solve the following optimization problem

$$\begin{array}{ll} \max & Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} \\ \text{subject to} & \|\mathbf{x}\|^2 = 1 \end{array}$$

### Solution

The characteristic equation is

$$|A - \lambda I| = 0 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

The maximum eigenvalue is  $\lambda = 6$  and its corresponding eigenvector is

$$\mathbf{u} = \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right). \text{ Therefore, the maximum of } Q(\mathbf{x}) \text{ is } 6 \text{ that is attained for}$$
$$\mathbf{x} = \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right).$$



## Theorem 3.2

Let  $A$ ,  $\lambda_{\max}$  and  $\mathbf{u}_{\max}$  be defined as in the previous theorem. Then the solution of

$$\begin{array}{ll} \max & Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} \\ \text{subject to} & \|\mathbf{x}\|^2 = 1 \\ & \mathbf{x} \cdot \mathbf{u}_{\max} = 0 \end{array}$$

is given by the second largest eigenvalue  $\lambda_{\max-1}$  that is attained for its associated eigenvector ( $\mathbf{u}_{\max-1}$ ).

## Exercises

From Lay (3rd ed.), Chapter 7, Section 3:

- 7.3.1
- 7.3.3
- 7.3.13

- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - Constrained optimization (b)
  - Singular Value Decomposition (SVD) (c)

# Singular Value Decomposition (SVD)

## Introduction

Unfortunately, not all matrices can be diagonalized and factorized as

$$A = PDP^{-1}$$

However, all of them (even rectangular matrices) can be factorized as

$$A = QDP^{-1}$$

This is called the **Singular Value Decomposition**. It imitates the property of stretching/shrinking of eigenvalues and eigenvectors. For instance, assume  $\mathbf{u}$  is an eigenvector, then

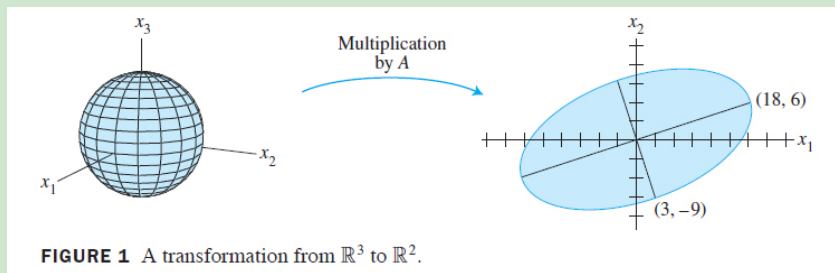
$$A\mathbf{u} = \lambda\mathbf{u} \Rightarrow \|A\mathbf{u}\| = |\lambda|\|\mathbf{u}\|$$

If  $|\lambda| > 1$ , then the transformed vector  $A\mathbf{u}$  is stretched with respect to  $\mathbf{u}$ . On the contrary, if  $|\lambda| < 1$ , then the transformed vector  $A\mathbf{u}$  is shrunk with respect to  $\mathbf{u}$ .

# Singular Value Decomposition (SVD)

## Example

Consider  $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$  and the linear transformation  $T(\mathbf{x}) = A\mathbf{x}$ . It transforms the unit sphere in  $\mathbb{R}^3$  onto an ellipse of  $\mathbb{R}^2$



Look for the direction that maximizes  $\|A\mathbf{x}\|$  subject to  $\|\mathbf{x}\| = 1$ .

# Singular Value Decomposition (SVD)

## Solution

We may maximize  $\|A\mathbf{x}\|^2$  because  $\|A\mathbf{x}\|$  is maximum iff  $\|A\mathbf{x}\|^2$  is maximum.

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^T A^T A \mathbf{x}$$

which is a quadratic form since  $A^T A$  is symmetric:

$$A^T A = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

By Theorem 3.1, the maximum eigenvalue is  $\max \|A\mathbf{x}\|^2 = \lambda_{\max} = 360$  and its associated eigenvector  $\mathbf{u}_{\max} = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$ . Consequently  $\max \|A\mathbf{x}\| = \sqrt{360} = 6\sqrt{10}$  that is attained for

$$A\mathbf{u}_{\max} = \begin{pmatrix} 18 \\ 6 \end{pmatrix}$$

# Singular Value Decomposition (SVD)

## Definition 4.1

*Singular Values of a matrix* Let  $A \in \mathcal{M}_{m \times n}$ .  $A^T A$  can always be orthogonally diagonalized. Let  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  a base of  $\mathbb{R}^n$  formed by the eigenvectors of  $A^T A$  and let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be its corresponding eigenvalues. Then

$$\|A\mathbf{v}_i\|^2 = (A\mathbf{v}_i)^T (A\mathbf{v}_i) = \mathbf{v}_i^T A^T A \mathbf{v}_i = \mathbf{v}_i^T (\lambda_i \mathbf{v}_i) = \lambda_i \|\mathbf{v}_i\|^2$$

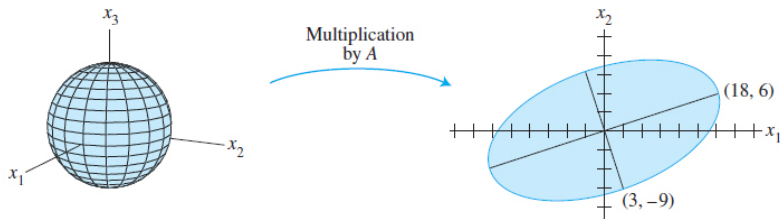
If we take the square root

$$\|A\mathbf{v}_i\| = \sqrt{\lambda_i} \|\mathbf{v}_i\|$$

That is,  $\sqrt{\lambda_i}$  reflects the amount by which  $\mathbf{v}_i$  is stretched or shrinked.  $\sqrt{\lambda_i}$  is called a **singular value** and it is denoted as  $\sigma_i$ .

# Singular Value Decomposition (SVD)

## Example (continued)



**FIGURE 1** A transformation from  $\mathbb{R}^3$  to  $\mathbb{R}^2$ .

In the example of Slide 45, the singular values are the lengths of the ellipse in  $\mathbb{R}^2$  and they are  $6\sqrt{10}$ ,  $3\sqrt{10}$  and 0. From the singular values we learn that the unit sphere in  $\mathbb{R}^3$  (there are 3 singular values) is collapsed in 2D (one of the singular values is 0) onto an ellipse (the remaining two singular values are different from each other).



# Singular Value Decomposition (SVD)

## Theorem 4.1

Let  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  a basis of  $\mathbb{R}^n$  formed by the eigenvectors of  $A^T A$  sorted in descending order and let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be its corresponding eigenvalues. Let us assume that  $A$  has  $r$  non-null singular values. Then

$$S = \{\mathbf{A}\mathbf{v}_1, \mathbf{A}\mathbf{v}_2, \dots, \mathbf{A}\mathbf{v}_r\}$$

is a basis of  $\text{Col}\{A\}$  and

$$\text{Rank}\{A\} = r$$

### Proof

By Theorem 1.1, any two eigenvectors are orthogonal to each other if they correspond to different eigenvalues, that is,  $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ . Then,

$$(\mathbf{A}\mathbf{v}_i) \cdot (\mathbf{A}\mathbf{v}_j) = \mathbf{v}_i^T A^T \mathbf{A} \mathbf{v}_j = \mathbf{v}_i^T (\lambda_j \mathbf{v}_j) = \lambda_j (\mathbf{v}_i^T \mathbf{v}_j) = \lambda_j (\mathbf{v}_i \cdot \mathbf{v}_j) = 0$$

That is  $\mathbf{A}\mathbf{v}_i$  and  $\mathbf{A}\mathbf{v}_j$  are also orthogonal.

# Singular Value Decomposition (SVD)

Additionally, if the eigenvectors  $\mathbf{v}_i$  are unitary, then (see Definition 4.1)

$$\sigma_i = \|\mathbf{A}\mathbf{v}_i\|$$

Since there are  $r$  non-null singular values,  $\mathbf{A}\mathbf{v}_i \neq \mathbf{0}$  only for  $i = 1, 2, \dots, r$ . So the set  $S$  is a set of non-null, orthogonal vectors. To show it is a basis of  $\text{Col}\{A\}$  we still need to show that any vector in  $\text{Col}\{A\}$  can be expressed as a linear combination of the vectors in  $S$ . We know that the eigenvalues of  $A^T A$  is a basis of  $\mathbb{R}^n$ . Then for any vector  $\mathbf{x} \in \mathbb{R}^n$  there exist coefficients  $c_1, c_2, \dots, c_n$  not all of them zero such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$$

If we transform this vector

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{A}(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n) && \text{[Linear transformation]} \\ &= c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2 + \dots + c_n\mathbf{A}\mathbf{v}_n && \text{[non-null singular values]} \\ &= c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2 + \dots + c_r\mathbf{A}\mathbf{v}_r \end{aligned}$$

# Singular Value Decomposition (SVD)

That is any transformed vector  $Ax$  can be expressed as a linear combination of the elements in  $S$ . Consequently,  $S$  is a basis of  $\text{Col}\{A\}$ .

Finally,  $\text{Rank}\{A\}$  is nothing more than the dimension of  $\text{Col}\{A\}$ . Since  $A$  is a basis of  $\text{Col}\{A\}$  and it has  $r$  vectors, then  $\text{Rank}\{A\} = r$ .

# Singular Value Decomposition (SVD)

## Theorem 4.2 (The Singular Value Decomposition)

Let  $A \in \mathcal{M}_{m \times n}$  be a matrix with rank  $r$ . Then, there exists a matrix  $\Sigma \in \mathcal{M}_{m \times n}$  whose diagonal entries are the first  $r$  singular values of  $A$  sorted in descending order ( $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ ) and there exist orthogonal matrices  $U \in \mathcal{M}_{m \times m}$  and  $V \in \mathcal{M}_{n \times n}$  such that

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

$\Sigma$  is unique but  $U$  and  $V$  are not. The columns of  $U$  are called the left singular vectors, and the columns of  $V$  are the right singular vectors.

## Example

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} & v_{31} & v_{41} \\ v_{12} & v_{22} & v_{32} & v_{42} \\ v_{13} & v_{23} & v_{33} & v_{43} \\ v_{14} & v_{24} & v_{34} & v_{44} \end{pmatrix}$$

# Singular Value Decomposition (SVD)

## Proof

Let  $\lambda_i$  and  $\mathbf{v}_i$  ( $i = 1, 2, \dots, n$ ) be the eigenvalues and eigenvectors of  $A^T A$ . By Theorem 4.1 we know that  $S = \{\mathbf{A}\mathbf{v}_1, \mathbf{A}\mathbf{v}_2, \dots, \mathbf{A}\mathbf{v}_r\}$  is an orthogonal basis of  $\text{Col}\{A\}$ . Let's normalize these vectors

$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i} \quad i = 1, 2, \dots, r$$

and we extend the set  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$  to be an orthogonal basis of  $\mathbb{R}^m$ . Let us construct the matrices

$$U = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_m)$$
$$V = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n)$$

By construction  $U$  and  $V$  are orthogonal, and

$$\begin{aligned} AV &= (\mathbf{A}\mathbf{v}_1 \quad \mathbf{A}\mathbf{v}_2 \quad \dots \quad \mathbf{A}\mathbf{v}_r \quad \mathbf{0} \quad \dots \quad \mathbf{0}) \\ &= (\sigma_1 \mathbf{u}_1 \quad \sigma_2 \mathbf{u}_2 \quad \dots \quad \sigma_r \mathbf{u}_r \quad \mathbf{0} \quad \dots \quad \mathbf{0}) \end{aligned}$$

# Singular Value Decomposition (SVD)

Proof (continued)

On the other side, let

$$D = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_r \end{pmatrix} \quad \Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

Then,

$$U\Sigma = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_m) \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = (\sigma_1\mathbf{u}_1 \quad \sigma_2\mathbf{u}_2 \quad \dots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \dots \quad \mathbf{0})$$

Therefore,

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

since  $V$  is orthogonal.

# Singular Value Decomposition (SVD)

## Theorem 4.3 (Properties of the SVD decomposition)

*In a SVD decomposition*

- *The left singular vectors of  $A$  are eigenvectors of  $AA^T$ .*
- *The right singular vectors of  $A$  are eigenvectors of  $A^T A$ .*
- *The singular values are the square root of the eigenvalues of both  $AA^T$  and  $A^T A$ .*
- *The singular values are the length of the semiaxes of the mapping of the unit hypersphere in  $\mathbb{R}^n$  onto  $\mathbb{R}^m$ .*
- *The columns of  $U$  form an orthogonal basis of  $\mathbb{R}^m$ .*
- *The columns of  $V$  form an orthogonal basis of  $\mathbb{R}^n$ .*

# Singular Value Decomposition (SVD)

## Example

Let's calculate the SVD decomposition of  $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$ .

Step 1: Orthogonally diagonalize  $A^T A$

$$A^T A = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

Its eigenvalues and eigenvectors are

$$\begin{aligned} \lambda_1 &= 360 & \mathbf{v}_1 &= \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) \\ \lambda_2 &= 90 & \mathbf{v}_2 &= \left(-\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right) \\ \lambda_3 &= 0 & \mathbf{v}_3 &= \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right) \end{aligned}$$



# Singular Value Decomposition (SVD)

Step 2: Construct  $V$  and  $\Sigma$

$$V = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3) = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$
$$\Sigma = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \end{pmatrix} = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}$$

Step 3: Construct  $U$

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1} = \left( \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right)$$
$$\mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2} = \left( \frac{1}{\sqrt{10}}, -\frac{3}{\sqrt{10}} \right)$$

The set  $\{\mathbf{u}_1, \mathbf{u}_2\}$  is already a basis of  $\mathbb{R}^2$ , so there is no need to extend it.

# Singular Value Decomposition (SVD)

Finally we have

$$\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$A = U\Sigma V^T$

MATLAB: `[U,S,V]=svd([4 11 14; 8 7 -2])`

# Singular Value Decomposition (SVD)

## Example

Let's calculate the SVD decomposition of  $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$ .

Step 1: Orthogonally diagonalize  $A^T A$

$$A^T A = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$$

Its eigenvalues and eigenvectors are

$$\begin{aligned} \lambda_1 &= 18 & \mathbf{v}_1 &= \left( \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \\ \lambda_2 &= 0 & \mathbf{v}_2 &= \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \end{aligned}$$

# Singular Value Decomposition (SVD)

Step 2: Construct  $V$  and  $\Sigma$

$$V = (\mathbf{v}_1 \quad \mathbf{v}_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$
$$\Sigma = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Step 3: Construct  $U$

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

The set  $\{\mathbf{u}_1\}$  is not yet a basis of  $\mathbb{R}^3$ , so we need to extend it with orthogonal vectors. All vectors orthogonal to  $\mathbf{u}_1$  fulfill

$$\mathbf{u}_1 \cdot \mathbf{u} = 0 = \frac{1}{3}x_1 - \frac{2}{3}x_2 + \frac{2}{3}x_3 \Rightarrow x_1 = 2x_2 - 2x_3$$

# Singular Value Decomposition (SVD)

Step 3: Construct  $U$  (continued)

A basis of this space is  $\mathbf{w}_2 = (2, 1, 0)$  and  $\mathbf{w}_3 = (-2, 0, 1)$ . But this basis is not orthogonal. Let's make it orthogonal following Gram-Schmidt procedure

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left( \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0 \right)$$

$$\mathbf{w}'_3 = \mathbf{w}_3 - \langle \mathbf{w}_3, \mathbf{v}_2 \rangle \mathbf{v}_2 = \left( -\frac{2}{5}, \frac{4}{5}, 1 \right)$$

$$\mathbf{u}_3 = \frac{\mathbf{w}'_3}{\|\mathbf{w}'_3\|} = \left( -\frac{2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{\sqrt{5}}{3} \right)$$

In fact, SVD does not require the  $\mathbf{u}$  vectors to be unitary, but it is simply convenient. We can make  $\mathbf{u}_2$  and  $\mathbf{u}_3$  unitary because they are “free” (we are constructing them simply to extend the set of  $\mathbf{u}$  vectors to be a basis of  $\mathbb{R}^3$ ), but not  $\mathbf{u}_1$  because it is “bound” to the singular value.

# Singular Value Decomposition (SVD)

Finally we have

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{3\sqrt{5}}{4} \\ \frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$A = U\Sigma V^T$

# Algebraic applications of SVD

## Matrix condition number

Let  $\sigma_1$  and  $\sigma_r$  be the largest and smallest singular values of a matrix  $A$ . The condition number of the matrix is defined as

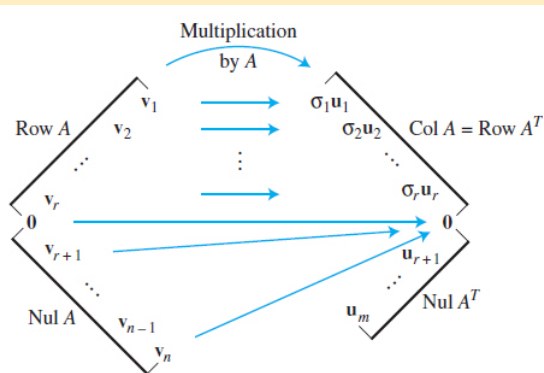
$$\kappa(A) = \frac{\sigma_1}{\sigma_r}$$

If this condition number is very large, the equations system  $A\mathbf{x} = \mathbf{b}$  is ill-posed and small perturbations in  $\mathbf{b}$  translate into large perturbations in  $\mathbf{x}$ . As a rule of thumb, if  $\kappa(A) = 10^k$ , then you may lose up to  $k$  digits of accuracy.

# Algebraic applications of SVD

## Bases for fundamental spaces

The  $U$  and  $V$  matrices provide bases for  $\text{Row}\{A\}$ ,  $\text{Col}\{A\} = \text{Row}\{A^T\}$ ,  $\text{Nul}\{A\}$  and  $\text{Nul}\{A^T\}$



**FIGURE 4** The four fundamental subspaces and the action of  $A$ .



## Theorem 4.4 (The Invertible Matrix Theorem (continued))

*The Invertible Matrix Theorem has been developed in Theorems 5.1 and 11.5 of Chapter 3, Theorem 10.5 of Chapter 5, Theorem 2.1 of Chapter 6. Here, we give an extension if  $A$  is invertible, then the following statements are equivalent to the previous statements:*

- xxvii.  $(\text{Col}\{A\})^\perp = \{\mathbf{0}\}$ .
- xxviii.  $(\text{Nul}\{A\})^\perp = \mathbb{R}^n$ .
- xxix.  $(\text{Row}\{A\}) = \mathbb{R}^n$ .
- xxx.  $A$  has  $n$  non-null singular values.

# Algebraic applications of SVD

## Reduced SVD and pseudoinverse of $A$

If within  $U$  and  $V$  we distinguish two submatrices, each one with  $r$  columns we have

$$U = (U_r U_{m-r}) \text{ and } V = (V_r V_{n-r})$$

Then,

$$A = U\Sigma V^T = (U_r U_{m-r}) \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_r^T \\ V_{n-r}^T \end{pmatrix} = U_r D V_r^T$$

Despite the fact that we may have removed many columns of  $U$  and  $V$ , we have not lost any information and the recovery of  $A$  is exact. The Moore-Penrose pseudoinverse is defined as

$$A^+ = V_r D^{-1} U_r^T$$

that is a  $n \times m$  matrix such that

$$A^+ A A^+ = A^+ \quad A A^+ A = A$$

# Algebraic applications of SVD

## Pseudoinverse of $A$ and Least Squares

It can be shown that the least-squares solution of the equation system  $A\mathbf{x} = \mathbf{b}$  is given by

$$\hat{\mathbf{x}} = A^+\mathbf{b}$$

## Matrix approximation

If instead of taking  $r$  components in the split of  $U$  and  $V$  (see previous slide) we take only  $k$  (assuming singular values have been ordered in descending order), and we reconstruct  $A_k$

$$A_k = U_k D_k V_k^T$$

This matrix is the matrix of rank  $k$  that minimizes the Frobenius norm of the difference

$$A_k = \min_{\text{Rank}\{B\}=k} \|A - B\|_F^2 = \min_{\text{Rank}\{B\}=k} \sum_{i,j=1}^n (a_{ij} - b_{ij})^2$$

## Exercises

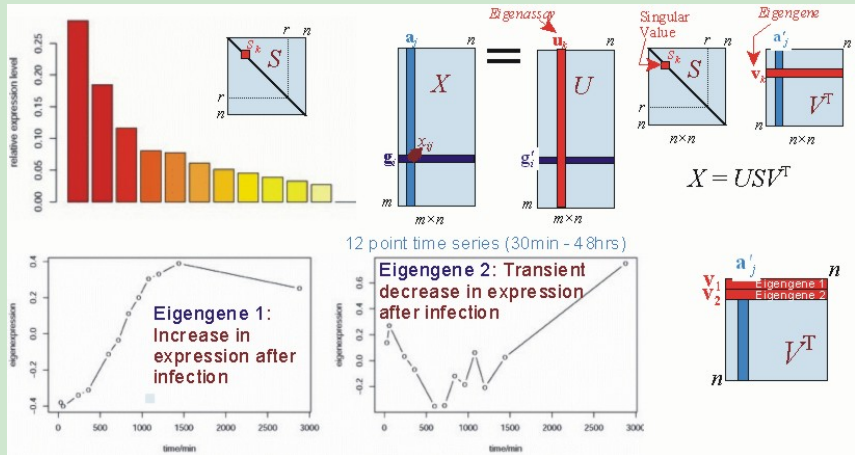
From Lay (3rd ed.), Chapter 7, Section 4:

- 7.4.3
- 7.4.11
- 7.4.15
- 7.4.17
- 7.4.18
- 7.4.19
- 7.4.20
- 7.4.23
- 7.4.24

# Applications of SVD

## Eigengenes and eigenassays

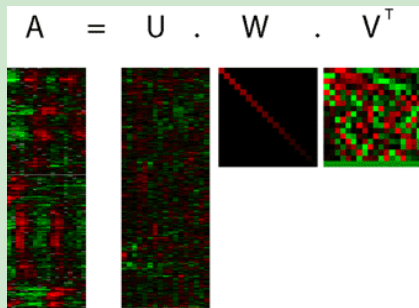
SVD is very much used to analyze the response of different genes to different assays or conditions.



# Applications of SVD

## Eigengenes and eigenassays

SVD is very much used to analyze the response of different genes to different assays or conditions.



Alter, O., Brown, P. O. and Botstein, D. (2000) Proc. Natl. Acad. Sci. USA 97, 10101

# Applications of SVD

## Eigenfaces

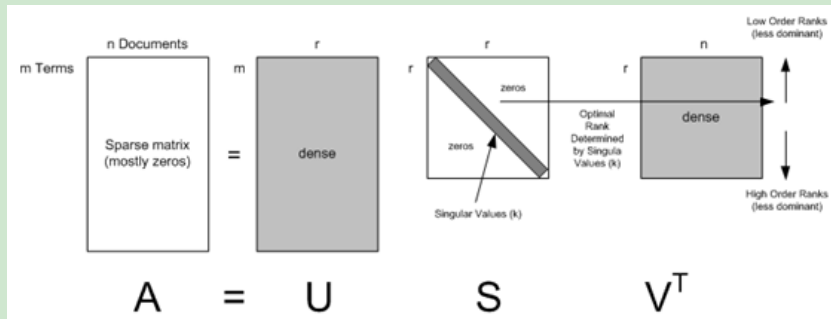
In this example we see the effect of matrix approximation by the reduced SVD.



# Applications of SVD

## Eigenfaces

We can also use SVD to automatically analyze documents.



P. Marksberry, D. Parsley. Managing the IE (Industrial Engineering) Mindset: A quantitative investigation of Toyota's practical thinking shared among employees. J. Industrial Engineering and Management, 4: 771-799 (2011)



- 8 Symmetric matrices and quadratic forms
  - Diagonalization of symmetric matrices (a)
  - Quadratic forms (b)
  - Constrained optimization (b)
  - Singular Value Decomposition (SVD) (c)