SYMMETRIC MATRICES AND QUADRATIC FORMS

DIAGONALIZATION OF SYMMETRIC MATRICES

DIAGONALIZATION OF SYMMETRIC MATRICES

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

DIAGONALIZATION OF SYMMETRIC MATRICES

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

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

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

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$$

DIAGONALIZATION OF SYMMETRIC MATRICES

An n* n matrix A is said to be **orthogonally diagonalizable if there** are an orthogonal matrix P and a diagonal matrix D such that

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

DIAGONALIZATION OF SYMMETRIC MATRICES

An n* n matrix A is said to be **orthogonally diagonalizable if there** are an orthogonal matrix P and a diagonal matrix D such that

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

Such a diagonalization requires n linearly independent and orthonormal eigenvectors. When is this possible?

DIAGONALIZATION OF SYMMETRIC MATRICES

An n* n matrix A is said to be **orthogonally diagonalizable if there** are an orthogonal matrix P and a diagonal matrix D such that

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

Such a diagonalization requires n linearly independent and orthonormal eigenvectors. When is this possible?

$$A^T = (PDP^T)^T = P^{TT}D^TP^T = PDP^T = A$$

DIAGONALIZATION OF SYMMETRIC MATRICES

An n* n matrix A is said to be **orthogonally diagonalizable if there** are an orthogonal matrix P and a diagonal matrix D such that

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

Such a diagonalization requires n linearly independent and orthonormal eigenvectors. When is this possible?

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A$$

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.

EXAMPLE 2 If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

EXAMPLE 2 If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

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

$$\lambda = 8: \ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6: \ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3: \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

EXAMPLE 2 If possible, diagonalize the matrix $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

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

$$\lambda = 8: \ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6: \ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3: \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

EXAMPLE 3 Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

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

EXAMPLE 3 Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

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

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

EXAMPLE 3 Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, whose characteristic equation is

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

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix} \qquad \mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|}2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

The set of eigenvalues of a matrix A is sometimes called the *spectrum* of A

The Spectral Theorem for Symmetric Matrices

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

a. 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. A is orthogonally diagonalizable.

Spectral Decomposition

spectral decomposition of A

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Spectral Decomposition

spectral decomposition of A

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \mathbf{u}_1 & \cdots & \lambda_n \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Spectral Decomposition

spectral decomposition of A

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \mathbf{u}_1 & \cdots & \lambda_n \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

an $n \times n$ matrix of rank 1

QUADRATIC FORM

Quadratic form

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

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

Change of Variable in a Quadratic Form

$$\mathbf{x} = P\mathbf{y}, \quad \mathbf{y} = P^{-1}\mathbf{x}$$

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

$$\mathbf{y}^T D\mathbf{y}$$

there is an *orthogonal* matrix P such that $P^T AP$ is a diagonal matrix D

EXAMPLE 1 Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^T A \mathbf{x}$ for the following matrices:

a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

EXAMPLE 1 Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Compute $\mathbf{x}^T A \mathbf{x}$ for the following matrices:

a. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$ b. $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

a. $\mathbf{x}^T A \mathbf{x} = [\,x_1 \quad x_2\,] \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,] \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2.$

b. There are two -2 entries in A. Watch how they enter the calculations. The $(1, 2)$-entry in A is in boldface type.

$$\begin{aligned}
\mathbf{x}^T A \mathbf{x} &= [\,x_1 \quad x_2\,] \begin{bmatrix} 3 & \mathbf{-2} \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\,x_1 \quad x_2\,] \begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\
&= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\
&= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\
&= 3x_1^2 - 4x_1x_2 + 7x_2^2
\end{aligned}$$

EXAMPLE 4 Make a change of variable that transforms the quadratic form in Example 3 into a quadratic form with no cross-product term.

SOLUTION The matrix of the quadratic form in Example 3 is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

EXAMPLE 4 Make a change of variable that transforms the quadratic form in Example 3 into a quadratic form with no cross-product term.

SOLUTION The matrix of the quadratic form in Example 3 is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

The first step is to orthogonally diagonalize A. Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$. Associated unit eigenvectors are

$$\lambda = 3\colon \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7\colon \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

EXAMPLE 4 Make a change of variable that transforms the quadratic form in Example 3 into a quadratic form with no cross-product term.

SOLUTION The matrix of the quadratic form in Example 3 is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

The first step is to orthogonally diagonalize A. Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = -7$. Associated unit eigenvectors are

$$\lambda = 3: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) \\ &= \mathbf{y}^T P^T A P\mathbf{y} = \mathbf{y}^T D\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned}$$

Quadratic form

The Principal Axes Theorem

Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

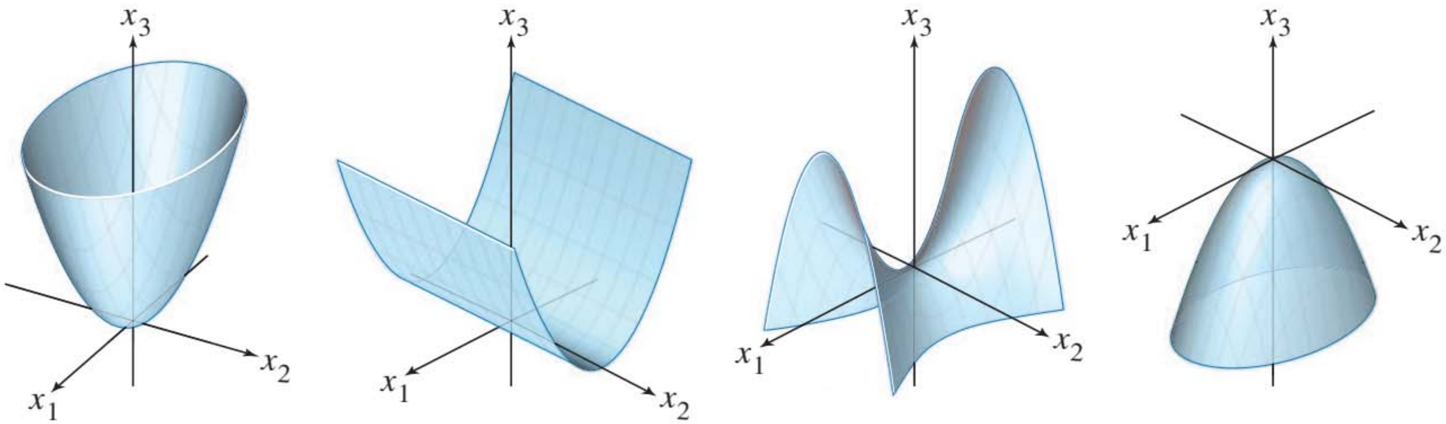
Quadratic form

The Principal Axes Theorem

Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.

The columns of P in the theorem are called the **principal axes** of the quadratic form $\mathbf{x}^T A\mathbf{x}$.

Classifying quadratic forms



(a) $z = 3x_1^2 + 7x_2^2$ (b) $z = 3x_1^2$ (c) $z = 3x_1^2 - 7x_2^2$ (d) $z = -3x_1^2 - 7x_2^2$

FIGURE 4 Graphs of quadratic forms.

$$(x_1, x_2, z) \text{ where } z = Q(\mathbf{x})$$

Classifying Quadratic Forms

A quadratic form Q is:

a. **positive definite** if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$,

b. **negative definite** if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$,

c. **indefinite** if $Q(\mathbf{x})$ assumes both positive and negative values.

Also, Q is said to be **positive semidefinite** if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, and to be **negative semidefinite** if $Q(\mathbf{x}) \leq 0$ for all $\mathbf{x}$.

Classifying Quadratic Forms

Quadratic Forms and Eigenvalues

Let A be an $n \times n$ symmetric matrix. Then a quadratic form $\mathbf{x}^T A\mathbf{x}$ is:

a. positive definite if and only if the eigenvalues of A are all positive,

b. negative definite if and only if the eigenvalues of A are all negative, or

c. indefinite if and only if A has both positive and negative eigenvalues.

Proof

PROOF By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$ such that

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

positive definite matrix A is a *symmetric* matrix for which the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite

CONSTRAINED OPTIMIZATION

CONSTRAINED OPTIMIZATION

Engineers, economists, scientists, and mathematicians often need to find the maximum or minimum value of a quadratic form Q(x) for **x** in some specified set.

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1$$

CONSTRAINED OPTIMIZATION

Engineers, economists, scientists, and mathematicians often need to find the maximum or minimum value of a quadratic form Q(x) for **x** in some specified set.

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1$$

When a quadratic form Q has no cross-product terms, it is easy to find the maximum and minimum of $Q(\mathbf{x})$ for $\mathbf{x}^T\mathbf{x} = 1$.

example

EXAMPLE 1 Find the maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

example

EXAMPLE 1 Find the maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\leq 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

$$\mathbf{x} = (1, 0, 0)$$

example

EXAMPLE 1 Find the maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

$$
\begin{aligned}
Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\
&\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\
&= 9(x_1^2 + x_2^2 + x_3^2) \\
&= 9
\end{aligned}
$$

$$\mathbf{x} = (1, 0, 0)$$

$$Q(\mathbf{x}) \ge 3x_1^2 + 3x_2^2 + 3x_3^2 = 3(x_1^2 + x_2^2 + x_3^2) = 3$$

CONSTRAINED OPTIMIZATION

$$m = \min \{\mathbf{x}^T A \mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \max \{\mathbf{x}^T A \mathbf{x} : \|\mathbf{x}\| = 1\}$$

CONSTRAINED OPTIMIZATION

$$m = \min\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}$$

Let A be a symmetric matrix, and define m and M as in (2). Then M is the greatest eigenvalue λ_1 of A and m is the least eigenvalue of A. The value of $\mathbf{x}^T A\mathbf{x}$ is M when $\mathbf{x}$ is a unit eigenvector $\mathbf{u}_1$ corresponding to M. The value of $\mathbf{x}^T A\mathbf{x}$ is m when $\mathbf{x}$ is a unit eigenvector corresponding to m.

CONSTRAINED OPTIMIZATION

$$m = \min\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}$$

Let A be a symmetric matrix, and define m and M as in (2). Then M is the greatest eigenvalue λ_1 of A and m is the least eigenvalue of A. The value of $\mathbf{x}^T A\mathbf{x}$ is M when $\mathbf{x}$ is a unit eigenvector $\mathbf{u}_1$ corresponding to M. The value of $\mathbf{x}^T A\mathbf{x}$ is m when $\mathbf{x}$ is a unit eigenvector corresponding to m.

PROOF Orthogonally diagonalize A as PDP^{-1}. We know that

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} \quad \text{when } \mathbf{x} = P\mathbf{y}$$

$$\|\mathbf{x}\| = \|P\mathbf{y}\| = \|\mathbf{y}\| \quad \text{for all } \mathbf{y}$$

CONSTRAINED OPTIMIZATION

To simplify notation, suppose that A is a 3×3 matrix with eigenvalues $a \geq b \geq c$.

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

$$\begin{aligned} \mathbf{y}^T D\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

CONSTRAINED OPTIMIZATION

To simplify notation, suppose that A is a 3×3 matrix with eigenvalues $a \geq b \geq c$.

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \qquad \begin{aligned} \mathbf{y}^T D\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

Thus $M \leq a$, by definition of M. However, $\mathbf{y}^T D\mathbf{y} = a$ when $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$

$M = a$

CONSTRAINED OPTIMIZATION

To simplify notation, suppose that A is a 3×3 matrix with eigenvalues $a \geq b \geq c$.

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

$$\begin{aligned} \mathbf{y}^T D\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

Thus $M \leq a$, by definition of M. However, $\mathbf{y}^T D\mathbf{y} = a$ when $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$

$M = a$

$$\mathbf{x} = P\mathbf{e}_1 = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_1$$

EXAMPLE 3 Let $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Find the maximum value of the quadratic form $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which this maximum value is attained.

EXAMPLE 3 Let $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Find the maximum value of the quadratic form $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which this maximum value is attained.

$$0 = -\lambda^3 + 10\lambda^2 - 27\lambda + 18 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \text{ Set } \mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}.$$

CONSTRAINED OPTIMIZATION

Let A, λ_1, and $\mathbf{u}_1$ be as in Theorem 6. Then the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue, λ_2, and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to λ_2.

CONSTRAINED OPTIMIZATION

Let A, λ_1, and $\mathbf{u}_1$ be as in Theorem 6. Then the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue, λ_2, and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to λ_2.

Let A be a symmetric $n \times n$ matrix with an orthogonal diagonalization $A = PDP^{-1}$, where the entries on the diagonal of D are arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and where the columns of P are corresponding unit eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then for $k = 2, \ldots, n$, the maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0, \quad \ldots, \quad \mathbf{x}^T\mathbf{u}_{k-1} = 0$$

is the eigenvalue λ_k, and this maximum is attained at $\mathbf{x} = \mathbf{u}_k$.

SINGULAR VALUE DECOMPOSITION

Introduction

The absolute values of the eigenvalues of a symmetric matrix A measure the amounts that A stretches or shrinks certain vectors (the eigenvectors)

$$\|A\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| = |\lambda|$$

Introduction

The absolute values of the eigenvalues of a symmetric matrix A measure the amounts that A stretches or shrinks certain vectors (the eigenvectors)

$$\|A\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| = |\lambda|$$

Example

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

$\mathbf{x}$ at which the length $\|A\mathbf{x}\|$ is maximized, and compute this maximum length $\|\mathbf{x}\| = 1$

Introduction

The absolute values of the eigenvalues of a symmetric matrix A measure the amounts that A stretches or shrinks certain vectors (the eigenvectors)

$$\|A\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| = |\lambda|$$

Example

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

$\mathbf{x}$ at which the length $\|A\mathbf{x}\|$ is maximized, and compute this maximum length $\|\mathbf{x}\| = 1$

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

$\boxed{\text{the greatest eigenvalue } \lambda_1 \text{ of } A^T A}$

example

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix} \qquad \lambda_1 = 360, \lambda_2 = 90, \text{ and } \lambda_3 = 0$$

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

For $\|\mathbf{x}\| = 1$, the maximum value of $\|A\mathbf{x}\|$ is $\|A\mathbf{v}_1\| = \sqrt{360} = 6\sqrt{10}$.

Singular Values

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

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

Singular Values

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

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

eigenvalues of A^TA are all nonnegative

The **singular values** of A are the square roots of the eigenvalues of A^TA, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order. That is, $\sigma_i = \sqrt{\lambda_i}$ for $1 \leq i \leq n$.

Singular Values

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

$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i \\ &= \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) && \text{Since } \mathbf{v}_i \text{ is an eigenvector of } A^TA \\ &= \lambda_i && \text{Since } \mathbf{v}_i \text{ is a unit vector} \end{aligned}$$

eigenvalues of A^TA are all nonnegative

The **singular values** of A are the square roots of the eigenvalues of A^TA, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order. That is, $\sigma_i = \sqrt{\lambda_i}$ for $1 \le i \le n$.

Singular Values

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

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$

$$
\begin{aligned}
\|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i \\
&= \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) && \text{Since } \mathbf{v}_i \text{ is an eigenvector of } A^TA \\
&= \lambda_i && \text{Since } \mathbf{v}_i \text{ is a unit vector}
\end{aligned}
$$

eigenvalues of A^TA are all nonnegative

The **singular values** of A are the square roots of the eigenvalues of A^TA, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order. That is, $\sigma_i = \sqrt{\lambda_i}$ for $1 \leq i \leq n$.

the singular values of A are the lengths of the vectors $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$

Theorem

THEOREM 9

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

Theorem

THEOREM 9 Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A^TA, arranged so that the corresponding eigenvalues of A^TA satisfy $\lambda_1 \geq \cdots \geq \lambda_n$, and suppose A has r nonzero singular values. Then $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is an orthogonal basis for Col A, and rank $A = r$.

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

Theorem

THEOREM 9

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

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

Col A

$$\mathbf{y} = A\mathbf{x} \qquad \mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

$$\begin{aligned}\mathbf{y} = A\mathbf{x} &= c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + c_{r+1}A\mathbf{v}_{r+1} + \cdots + c_n A\mathbf{v}_n \\ &= c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + 0 + \cdots + 0\end{aligned}$$

SVD

THEOREM 10 **The Singular Value Decomposition**

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

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

$$\Sigma = \begin{array}{ll} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} & \leftarrow m - r \text{ rows} \\ \qquad\quad \uparrow n - r \text{ columns} & \end{array}$$

SVD

THEOREM 10 **The Singular Value Decomposition**

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

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

$$\Sigma = \begin{array}{cl} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} & \leftarrow m - r \text{ rows} \\ \quad\uparrow n - r \text{ columns} & \end{array}$$

matrices U and V are not uniquely determined by A, but the diagonal entries of D are necessarily the singular values of A

columns of U **left singular vectors of A**
columns of V **right singular vectors of A**

proof

PROOF Let λ_i and $\mathbf{v}_i$ be as in Theorem 9, so that $\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$ is an orthogonal basis for Col A. Normalize each $A\mathbf{v}_i$ to obtain an orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$, where

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|}A\mathbf{v}_i = \frac{1}{\sigma_i}A\mathbf{v}_i \qquad\qquad A\mathbf{v}_i = \sigma_i\mathbf{u}_i$$

extend $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ to an orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ of $\mathbb{R}^m$

$$U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,] \quad \text{and} \quad V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,]$$

$$AV = [\,A\mathbf{v}_1 \quad \cdots \quad A\mathbf{v}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] = [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,]$$

proof

$$AV = [\,A\mathbf{v}_1 \quad \cdots \quad A\mathbf{v}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] = [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,]$$

$$U\Sigma = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,] \left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & 0 & & 0 \end{array}\right] = [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] = AV$$

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

example

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

example

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

$$\boxed{V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}}$$

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}$$

example

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

$$\boxed{V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}}$$

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}$$

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

example

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

$$\boxed{V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}}$$

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}$$

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

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

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

$$\boxed{A = \underbrace{\begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}}_{\uparrow\; U} \underbrace{\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}}_{\uparrow\; \Sigma} \underbrace{\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}}_{\uparrow\; V^T} \quad \blacksquare}$$

example

Find a singular value decomposition of $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$ $\qquad A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$

eigenvalues of A^TA are 18 and 0

$\sigma_1 = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = 0$.

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

example

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

$$A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

eigenvalues of A^TA are 18 and 0

$\sigma_1 = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = 0$.

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

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$V = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

example

Find a singular value decomposition of $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$ $\quad A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$

eigenvalues of A^TA are 18 and 0

$\sigma_1 = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = 0$.

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

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad V = [\mathbf{v}_1 \quad \mathbf{v}_2] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$A\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix}, \quad A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{u}_1 = \frac{1}{3\sqrt{2}} A\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

example

$$\mathbf{u}_1^T \mathbf{x} = 0 \qquad x_1 - 2x_2 + 2x_3 = 0$$

example

$$\mathbf{u}_1^T\mathbf{x} = 0 \qquad x_1 - 2x_2 + 2x_3 = 0$$

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

example

$$\mathbf{u}_1^T \mathbf{x} = 0 \qquad x_1 - 2x_2 + 2x_3 = 0 \qquad \mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Gram-Schmidt

$$\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

APPLICATIONS

SVD for Image Compression

$$A = USV^T = \sum_{i=1}^{r} \sigma_i \mathrm{u}_i \mathrm{v}_i^T$$

$$\Downarrow$$

$$A_k = \sigma_1 \mathrm{u}_1 \mathrm{v}_1^T + \sigma_2 \mathrm{u}_2 \mathrm{v}_2^T + \cdots + \sigma_k \mathrm{u}_k \mathrm{v}_k^T$$

SVD for Image Compression

$$f_{ij} \quad \text{Where} \quad f_{ij} \equiv f(x_i, y_j)$$

Redundancy exists in Images
Size of images
Compression

$$A = USV^T = \sum_{i=1}^{r} \sigma_i \mathrm{u}_i \mathrm{v}_i^T$$

$$A_k = \sigma_1 \mathrm{u}_1 \mathrm{v}_1{}^T + \sigma_2 \mathrm{u}_2 \mathrm{v}_2{}^T + \cdots + \sigma_k \mathrm{u}_k \mathrm{v}_k{}^T$$

The total storage for A_k will be

$$k\,(m + n + 1)$$

SVD for Image Compression

$$C_R = m*n/(k\,(m + n + 1))$$

To measure the quality between original image *A and the compressed image Ak, the* measurement of *Mean Square Error (MSE)*

$$MSE = \frac{1}{mn}\sum_{y=1}^{m}\sum_{x=1}^{n}(f_A(x,y) - f_{A_k}(x,y))$$

C_R Comp	MSE (Quality)
5.03	108.11
3.35	63.15
2.51	40.39
2.01	27.22
1.68	15.64
1.26	9.07
1	

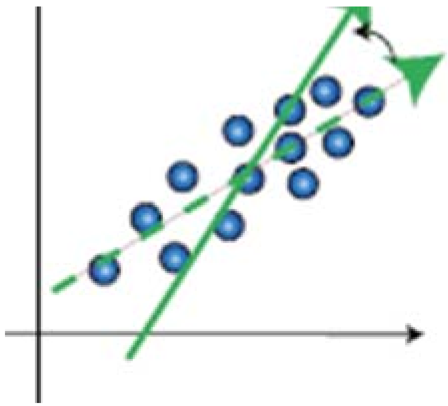
Face Recognition: PCA (principle component analysis)

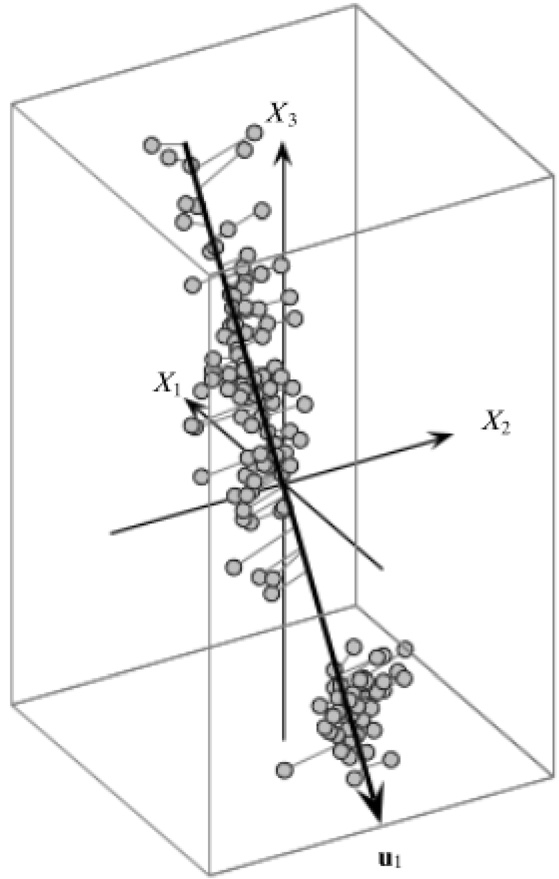
- Assume each face image has $m \times n = d$ *pixels*
- an $d \times 1$ column vector ***xi***
- A training set, *D with n* number of face images of known individuals forms an $d \times n$ *matrix:*

$$\mathbf{D} = \begin{array}{c|cccc} & X_1 & X_2 & \cdots & X_d \\ \hline \mathbf{x}_1^T & x_{11} & x_{12} & \cdots & x_{1d} \\ \mathbf{x}_2^T & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_n^T & x_{n1} & x_{n2} & \cdots & x_{nd} \end{array}$$

$$\overline{\mathbf{D}} = \mathbf{D} - \mathbf{1} \cdot \boldsymbol{\mu}^T$$

PRINCIPAL COMPONENT ANALYSIS




X_3
X_1
X_2
$\mathbf{u}_1$

Best Line Approximation: We will start with $r = 1$, that is, the one-dimensional subspace or line u that best approximates D in terms of the variance of the projected points

PCA

Best Line Approximation: We will start with r = 1, that is, the one-dimensional subspace or line u that best approximates D in terms of the variance of the projected points

$$\boxed{\|\mathbf{u}\|^2 = \mathbf{u}^T\mathbf{u} = 1}$$

The projection of the centered point $\bar{\mathbf{x}}_i \in \overline{\mathbf{D}}$ on the vector $\mathbf{u}$

$$\mathbf{x}'_i = \left(\frac{\mathbf{u}^T\bar{\mathbf{x}}_i}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u} = (\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u} = a_i\mathbf{u}$$

choose the direction u such that the variance of the projected points is maximized

PCA

Best Line Approximation: We will start with r = 1, that is, the one-dimensional subspace or line u that best approximates D in terms of the variance of the projected points

$$\boxed{\|\mathbf{u}\|^2 = \mathbf{u}^T\mathbf{u} = 1}$$

The projection of the centered point $\bar{\mathbf{x}}_i \in \overline{\mathbf{D}}$ on the vector $\mathbf{u}$

$$\mathbf{x}_i' = \left(\frac{\mathbf{u}^T\bar{\mathbf{x}}_i}{\mathbf{u}^T\mathbf{u}}\right)\mathbf{u} = (\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u} = a_i\mathbf{u}$$

choose the direction u such that the variance of the projected points is maximized

$$\sigma_{\mathbf{u}}^2 = \frac{1}{n}\sum_{i=1}^{n}(a_i - \mu_a)^2 \qquad \mu_a = \frac{1}{n}\sum_{i=1}^{n}a_i = \frac{1}{n}\sum_{i=1}^{n}\mathbf{u}^T(\bar{\mathbf{x}}_i) = \mathbf{u}^T\bar{\boldsymbol{\mu}} = 0$$

PCA

$$\sigma_{\mathbf{u}}^2 = \frac{1}{n}\sum_{i=1}^{n}(a_i - \mu_a)^2 = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{u}^T\bar{\mathbf{x}}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{u}^T\left(\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T\right)\mathbf{u} = \mathbf{u}^T\left(\frac{1}{n}\sum_{i=1}^{n}\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T\right)\mathbf{u}$$

$$\boxed{\sigma_{\mathbf{u}}^2 = \mathbf{u}^T\mathbf{\Sigma}\mathbf{u}}$$

PCA

$$\sigma_{\mathbf{u}}^2 = \frac{1}{n}\sum_{i=1}^{n}(a_i - \mu_a)^2 = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{u}^T\bar{\mathbf{x}}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{u}^T\left(\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T\right)\mathbf{u} = \mathbf{u}^T\left(\frac{1}{n}\sum_{i=1}^{n}\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T\right)\mathbf{u}$$

$$\boxed{\sigma_{\mathbf{u}}^2 = \mathbf{u}^T\mathbf{\Sigma}\mathbf{u}}$$

where $\mathbf{\Sigma}$ is the sample covariance matrix for the centered data $\bar{\mathbf{D}}$

$$\max_{\mathbf{u}} \ \mathbf{u}^T\mathbf{\Sigma}\mathbf{u}$$
$$\mathbf{u}^T\mathbf{u} = 1$$

PCA

Minimum Squared Error Approach

direction that maximizes the projected variance is also the one that minimizes the average squared error

$$
\begin{aligned}
MSE(\mathbf{u}) &= \frac{1}{n}\sum_{i=1}^{n}\|\boldsymbol{\epsilon}_i\|^2 = \frac{1}{n}\sum_{i=1}^{n}\|\bar{\mathbf{x}}_i - \mathbf{x}'_i\|^2 = \frac{1}{n}\sum_{i=1}^{n}(\bar{\mathbf{x}}_i - \mathbf{x}'_i)^T(\bar{\mathbf{x}}_i - \mathbf{x}'_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\bar{\mathbf{x}}_i\|^2 - 2\bar{\mathbf{x}}_i^T\mathbf{x}'_i + (\mathbf{x}'_i)^T\mathbf{x}'_i\right) \qquad (7.15) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\bar{\mathbf{x}}_i\|^2 - 2\bar{\mathbf{x}}_i^T(\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u} + \left((\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u}\right)^T(\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u}\right), \text{since } \mathbf{x}'_i = (\mathbf{u}^T\bar{\mathbf{x}}_i)\mathbf{u} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\bar{\mathbf{x}}_i\|^2 - 2(\mathbf{u}^T\bar{\mathbf{x}}_i)(\bar{\mathbf{x}}_i^T\mathbf{u}) + (\mathbf{u}^T\bar{\mathbf{x}}_i)(\bar{\mathbf{x}}_i^T\mathbf{u})\mathbf{u}^T\mathbf{u}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\bar{\mathbf{x}}_i\|^2 - (\mathbf{u}^T\bar{\mathbf{x}}_i)(\bar{\mathbf{x}}_i^T\mathbf{u})\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\|\bar{\mathbf{x}}_i\|^2 - \frac{1}{n}\sum_{i=1}^{n}\mathbf{u}^T(\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T)\mathbf{u} \\
&= \frac{1}{n}\sum_{i=1}^{n}\|\bar{\mathbf{x}}_i\|^2 - \mathbf{u}^T\left(\frac{1}{n}\sum_{i=1}^{n}\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T\right)\mathbf{u}
\end{aligned}
$$

PCA

$$MSE = \sum_{i=1}^{n} \frac{\|\bar{\mathbf{x}}_i\|^2}{n} - \mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u}$$

$$\mathrm{var}(\mathbf{D}) = tr(\boldsymbol{\Sigma}) = \sum_{i=1}^{d} \sigma_i^2$$

$$MSE(\mathbf{u}) = \mathrm{var}(\mathbf{D}) - \mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u} = \sum_{i=1}^{d} \sigma_i^2 - \mathbf{u}^T \boldsymbol{\Sigma} \mathbf{u}$$

PCA

- Best 2-dimensional Approximation

- We already computed the direction with the most variance, namely u1, which is the eigenvector corresponding to the largest eigenvalue λ1

- We now want to find another direction v, which also maximizes the projected variance, but is orthogonal to u1.

PCA

- Best 2-dimensional Approximation
- We already computed the direction with the most variance, namely u1, which is the eigenvector corresponding to the largest eigenvalue λ1
- We now want to find another direction v, which also maximizes the projected variance, but is orthogonal to u1.

$$\max_{\mathbf{v}} \ \sigma_{\mathbf{v}}^2 = \mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v}$$

$$\mathbf{v}^T \mathbf{u}_1 = 0$$

$$\mathbf{v}^T \mathbf{v} = 1$$

second largest eigenvalue of °, with the second principal component being given by the corresponding eigenvector, that is, v = u_2.

PCA

Best $\boldsymbol{r}$-dimensional Approximation

To find the best r-dimensional approximation to $\boldsymbol{D}$, we compute the eigenvalues of Σ. Because Σ is positive semidefinite, its eigenvalues are non-negative

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r \geq \lambda_{r+1} \cdots \geq \lambda_d \geq 0$$

We select the r largest eigenvalues, and their corresponding eigenvectors to form the best r-dimensional approximation.

PCA

PCA (D, r)

$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i$ `// compute mean`
$\boldsymbol{Z} = \boldsymbol{D} - \boldsymbol{1} \cdot \boldsymbol{\mu}^T$ `// center the data`
$\Sigma = \frac{1}{n}\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)$ `// compute covariance matrix`

$(\lambda_1, \lambda_2, \ldots, \lambda_d) = \text{eigenvalues}(\Sigma)$ `// compute eigenvalues`
$\boldsymbol{U} = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_d\end{pmatrix} = \text{eigenvectors}(\Sigma)$ `// compute eigenvectors`

$\boldsymbol{U}_r = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_r\end{pmatrix}$ `// reduced basis`
$\boldsymbol{A} = \{\boldsymbol{a}_i \mid \boldsymbol{a}_i = \boldsymbol{U}_r^T \boldsymbol{x}_i, \text{for } i = 1, \ldots, n\}$ `// reduced dimensionality data`