Ch.3 Transformation

QSnake

An Introduction to Optimization

2/24,2023

Little Review

Definition inner product

The inner product is a real-valued function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ having the following properties:

1. Positivity: $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0$ if and only if $x = 0$
2. Symmetry: $\langle x, y \rangle = \langle y, x \rangle$
3. Additivity: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.
4. Homogeneity: $\langle rx, y \rangle = r \langle x, y \rangle$ for every $r \in \mathbb{R}$

Definition of Norm

The Euclidean norm of a vector $||x||$ has the following properties:

1. Positivity: $||x|| \geq 0, \; ||x|| = 0$ if and only if $x = 0$
2. Homogeneity: $||rx|| = |r|||x||, r \in \mathbb{R}$
3. Triangle inequality: $||x + y|| \leq ||x|| + ||y||$

Orthogonal

Let V be an inner product space. Vectors x and y in V are orthogonal if $\langle x, y \rangle = 0$. A subset S of V is orthogonal.

Review

A function $L : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if:

1. $L(ax) = aL(x)$ for every $x \in \mathbb{R}^n$ and $a \in \mathbb{R}$.
2. $L(x + y) = L(x) + L(y)$ for every $x, y \in \mathbb{R}^n$

and the linear transformation L can be represented by a matrix.(Friedberg Ch.2)

Adjoint Operator

$\langle Ax, x \rangle = \langle x, A^* x \rangle$
if A is a real matrix
$\langle Ax, x \rangle = \left\langle x, A^T \right\rangle$

Eigenvalues and Eigenvectors

Definition

Let A be an $n \times n$ real square matrix. A scalar λ and a non-zero vector v satisfying the equation $Av = \lambda v$ are said to be, respectively, an **eigenvalue** and an **eigenvector** of A.

Property

For λ to be an eigenvalue it is necessary and sufficient for the matrix $\lambda I - A$ to be singular, i.e. the **characteristic polynomial** of the matrix A equal 0.

Theorem

Suppose that the characteristic equation $\det[\lambda I - A] = 0$ has n distinct roots $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then, there exist n **linearly independent** vectors $v_1, \cdots, v_n$ such that

$$Av_i = \lambda_i v_i, \quad i = 1, 2, \cdots, n$$

Theorem

All eigenvalues of a real **symmetric** matrix are **real**.

Theorem

Any real symmetric $n \times n$ matrix has a set of n eigenvectors that are **mutually orthogonal**.

Orthogonal Projections

orthogonal complement

If V is a subspace of $\mathbb{R}^n$, then the orthogonal complement of V, denoted $V^\perp$, consists of all vectors that are orthogonal to every vector in V. Thus,

$$V^\perp = \{x : \langle v, Tx \rangle = 0 \forall v \in V\}$$

orthogonal decomposition

orthogonal projector

We say that a linear transformation P is an **orthogonal projector** onto V if for all $x \in \mathbb{R}^n$, we have $Px \in V$ and $x - Px \in V^\perp$

Example of orthogonal projector

we consider $\mathbb{R}^2$ here

Theorem

Let A be a given matrix.
Then, $R(A)^{\perp} = N(A^T)$ and $N(A)^{\perp} = R(A^T)$

Theorem

A matrix P is an orthogonal projector [onto the subspace $V = R(P)$] if and only if $P^2 = P = P^T$

Quadratic Forms

Quadratic Forms

A quadratic form $f : \mathbb{R}^n \to \mathbb{R}$ is a function

$$f(x) = x^T Qx =< x, Qx >$$

where Q is an $n \times n$ real matrix.

the generality of quadratic form

There is no loss of generality in assuming Q t be symmetric: $Q = Q^T$. For if the matrix Q is not symmetric, we can always replace it with the symmetric matrix

$$Q_0 = Q_0^T = \frac{1}{2}\left(Q + Q^T\right)$$

reason

terminology

1. positive definite: $x^T Qx > 0$ for all x non-zero vectors x.
2. positive semidefinite if $x^T Qx \geq 0$ for all x

and the negative is similar

principal minors and leading principal minor

Sylvester's Criterion

A quadraic form $x^T Qx,\ \ Q = Q^T$, is positive definite if and only if the leading principal minors of Q are positive.

Example

Consider $Q = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}$

Theorem 3.7

A symmetric matrix Q is positive definite (or positive semidefinite) if and only if all eigenvalue of Q are positive (or nonnegative).

Matrix Norms

Frobenius norm

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij})^2 \right)^{\frac{1}{2}}$$

where $A \in \mathbb{R}^{m \times n}$, and clearly it's a norm.

matrix induced norms

In many problems, both matrices and vectors appear simultaneously.

induced norms

We say that the matrix norm is induced by the given vector norms if for any matrix $A \in \mathbb{R}^{m \times n}$ and any vector $x \in \mathbb{R}^n$, the following inequality is satisfied:

$$||Ax||_{(m)} \leq ||A||||x||_{(n)}$$

We can define an induced matrix norm as

$$||A|| = \max_{||x||_{(n)}=1} ||Ax||_{(m)}$$

proof of matrix induced norms is norms

Theorem 3.8

Let

$$||x|| = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2} = \sqrt{\langle x, x\rangle}$$

The matrix norm induced by this vector norm is

$$||A|| = \sqrt{\lambda_1}$$

where λ_1 is the largest eigenvalue of the matrix $A^T A$

Rayleigh's Inequalities

If an $n \times n$ matrix P is real symmetric positive definite, then

$$\lambda_{\min}(P)||x||^2 \leq < x, Px > \leq \lambda_{\max}(P)||x||^2$$

where $\lambda_{\min}(P)$ denotes the smallest eigenvalue of P, and $\lambda_{\max}(P)$ denotes the largest eigenvalue of P.

Conclustion