Review: Linear Algebra

Danish (ddanish@cs.cmu.edu)
Office Hours : Wed 1:30 PM (outside GHC 8009)

01/26/2017

Slide/Content Courtesy : Dr. Zico Kolter

Overview

- Vector Space
- Matrix (Vector) properties and operations
  - Trace
  - Norms
  - Inverse
  - Rank (and linear independence)
  - Orthogonality
  - Eigenvalues and Eigenvectors
  - Quadratic Forms and Positive Semidefinite Matrices

The Trace

- $\mathrm{tr} : \mathbb{R}^{n \times n} \to \mathbb{R}$

$$\mathrm{tr}\, A = \sum_{i=1}^{n} A_{ii}$$

- Some properties
  - $\mathrm{tr}\, A = \mathrm{tr}\, A^T, \; A \in \mathbb{R}^{n \times n}$
  - $\mathrm{tr}(A + B) = \mathrm{tr}\, A + \mathrm{tr}\, B, \; A, B \in \mathbb{R}^{n \times n}$
  - $\boxed{\mathrm{tr}\, AB = \mathrm{tr}\, BA, \; A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}}$

Norms

- A vector norm is any function $f : \mathbb{R}^n \to \mathbb{R}$ with

  1. $f(x) \geq 0$ and $f(x) = 0 \Leftrightarrow x = 0$

  2. $f(ax) = |a| f(x)$ for $a \in \mathbb{R}$

  3. $f(x + y) \leq f(x) + f(y)$

- ℓ_2 norm

$$\|x\|_2 = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

- ℓ_1 norm

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

- ℓ_∞ norm

$$\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

Norms

- Geometric Interpretation
- Norms for matrices?

The Matrix Inverse

- Inverse of a square matrix $A \in \mathbb{R}^{n \times n}$ denoted A^{-1}

$$AA^{-1} = I = A^{-1}A$$

- May not exist (*non-singular* matrix has inverse, *singular* matrix does not)

$$A^{-1} \text{ exists } \iff Ax \neq 0 \text{ for all } x \neq 0$$

- Some important properties for $A, B \in \mathbb{R}^{n \times n}$ non-singular
  - $(A^{-1})^{-1} = A$
  - $(AB)^{-1} = B^{-1}A^{-1}$
  - $(A^T)^{-1} = (A^{-1})^T$

Solving Linear Equations

- Two linear equations

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

- In vector form, $Ax = b$, with

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} -13 \\ 9 \end{bmatrix}$$

- Solution using inverse

$$
\begin{aligned}
Ax &= b \\
A^{-1}Ax &= A^{-1}b \\
x &= A^{-1}b
\end{aligned}
$$

- Won't worry here about how to compute inverse, but it's very similar to the standard method for solving linear equations

Linear Independence and Rank

A set of vectors $x_1, x_2, ...x_n \subset \mathbb{R}^m$ are linearly independent if no vector can be represented as a linear combination of the remaining vectors. The rank of a matrix is the cardinality of the largest subset of the columns of some matrix A that is a linearly independent set.

Linear Independence and Rank

Linear Independence and Rank

- How to compute (row) Rank of a matrix?

Linear Independence and Rank

- How to compute (row) Rank of a matrix?
- Column Rank == Row Rank?

Linear Independence and Rank

- How to compute (row) Rank of a matrix?
- Column Rank == Row Rank?
- Maximum rank of a Matrix of size m x n?

Orthogonality

- Two vectors $x, y \in \mathbb{R}^n$ are *orthogonal* if

$$x^T y = 0$$

- They are *orthonormal* if, in addition,

$$\|x\|_2 = \|y\|_2 = 1$$

- A matrix $U \in \mathbb{R}^{n \times n}$ is orthogonal if all it's columns are orthonormal, i.e.,

$$U^T U = I = U U^T$$

- Columns of an orthogonal matrix are linearly independent

Eigenvalues and Eigenvectors

- For $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{C}$ is an *eigenvalue* and $x \in \mathbb{C}^n \neq 0$ an *eigenvector* if

$$Ax = \lambda x$$

- Satisfied if $(\lambda I - A)x = 0$, which we know exists if and only if $\det(\lambda I - A) = 0$

- $\det(\lambda I - A)$ is a polynomial (of degree n) in λ, its n roots are the n eigenvalues of A

Diagonalization

- Write equations for all n eigenvalues as

$$A \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

- Write as $AX = X\Lambda$, which implies

$$A = X \Lambda X^{-1}$$

if X is invertible (A *diagonalizable*)

- Important properties of eigenvectors/eigenvalues
  - $\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$
  - $\det A = \prod_{i=1}^{n} \lambda_i$
  - $\operatorname{rank}(A) =$ number of non-zero eigenvalues
  - Eigenvalues of A^{-1} are $1/\lambda_i$, $i = 1, \ldots, n$, eigenvectors are the same

Eigenvalues and Eigenvectors for Symmetric Matrices

- All eigenvalues are **real**
- The eigenvectors are orthonormal

Quadratic Forms

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

  $$f(x) = x^T A x$$

  for some $A \in \mathbb{R}^{n \times n}$

- Can take A to be symmetric, since

  $$x^T A x = (x^T A x)^T = x^T A^T x = x^T \frac{1}{2}(A + A^T) x$$

- $A \in \mathbb{R}^{n \times n}$ is positive definite (positive semidefinite) if $x^T Ax > 0$ ($x^T Ax \geq 0$) for all $x \in \mathbb{R}^n \neq 0$

- $A \in \mathbb{R}^{n \times n}$ is negative definite (negative semidefinite) if $x^T Ax < 0$ ($x^T Ax \leq 0$) for all $x \in \mathbb{R}^n \neq 0$

- A is indefinite if neither positive nor negative semidefinite

- Definiteness is characterized by eigenvalues of A
  - A positive definite $\Leftrightarrow \lambda_i > 0, \ \forall i$
  - A positive semidefinite $\Leftrightarrow \lambda_i \geq 0, \ \forall i$
  - A negative definite $\Leftrightarrow \lambda_i < 0, \ \forall i$
  - A negative semidefinite $\Leftrightarrow \lambda_i \leq 0, \ \forall i$

- Definiteness is characterized by eigenvalues of A
  - A positive definite $\Leftrightarrow \lambda_i > 0, \;\; \forall i$
  - A positive semidefinite $\Leftrightarrow \lambda_i \geq 0, \;\; \forall i$
  - A negative definite $\Leftrightarrow \lambda_i < 0, \;\; \forall i$
  - A negative semidefinite $\Leftrightarrow \lambda_i \leq 0, \;\; \forall i$

$$x^T A x = x^T U \Lambda U^T x = y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2$$