Review of Some Elements of Linear Algebra

Milan Vojnovic

ST445 Managing and Visualizing Data

Norms

- A norm is any function $f : \mathbf{R}^n \to \mathbf{R}_+$ that satisfies
  - $f(x) = 0 \Rightarrow x = 0$ (definite or point separating)
  - $f(x + y) \le f(x) + f(y)$ (subadditive or satisfying triangle inequality)
  - $f(\alpha x) = |\alpha| f(x)$ for all $\alpha \in \mathbf{R}$ (absolutely homogeneous)

- L_p norm: $\|x\|_p = (\sum_i |x_i|^p)^{1/p}, p \in \mathbf{R}, p \ge 1$

  - L_2 is referred as the Euclidean norm

- Max norm: $\|x\|_\infty = \max_i |x_i|$

- Frobenius norm: $\|A\|_F = \sqrt{\sum_{i,j} a_{i,j}^2}$ for a matrix A

Some special matrices

- A matrix is symmetric if it is equal to its transpose: $A = A^T$
- A unit vector is a vector with unit norm: $\|x\|_2 = 1$
- A vector x and a vector y are orthogonal if $x^T y = 0$
  - If, in addition, they are unit vectors, we say that they are orthonormal
- An orthogonal matrix is a square matrix whose rows (columns) are mutually orthonormal

$$A^T A = A A^T = I$$

- For any orthogonal matrix A the inverse matrix is equal to its transpose matrix:

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

Some special matrices (cont'd)

- A matrix whose eigenvalues are all strictly positive is said to be positive definite
- A matrix whose eigenvalues are all positive is said to be positive semidefinite
- Likewise, if all eigenvalues are negative or zero, the matrix is negative definite, and if all eigenvalues are negative, the matrix is negative semidefinite
- For any positive semidefinite matrix A: $\quad x^T A x \geq 0 \quad$ for all $\quad x$
  - In addition, if A is positive definite, then $x^T A x = 0 \quad \Rightarrow \quad x = 0$

Eigen-decomposition

- Suppose A is a $n \times n$ matrix with mutually orthogonal eigenvectors $v_1, v_2, \ldots, v_n$ and the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

- The eigen-decomposition of A is defined as:

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

where $V = (v_1, v_2, \ldots, v_n)$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

- Note every matrix has an eigen-decomposition

- Any real, symmetric matrix has an eigen-decomposition

Singular Value Decomposition (SVD)

- For an $m \times n$ matrix A the SVD is given by

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

where

U is an $m \times m$ matrix whose columns are left-singular vectors of A

Σ is an $m \times n$ diagonal matrix with the diagonal elements equal to the singular values of A

V is an $n \times n$ matrix whose columns are right-singular vectors of A

SVD (cont'd)

- The left-singular vectors of A correspond to the eigenvectors of AA^T

- The right-singular vectors of A correspond to the eigenvectors of $A^T A$

- The non-zero singular values of A are the square roots of the eigenvalues of $A^T A$

    - Same is true for AA^T

The Moore-Penrose pseudoinverse

- The pseudoinverse of A is defined as a matrix

$$A^+ = \lim_{\alpha \downarrow 0} (A^T A + \alpha I)^{-1} A^T$$

- Suppose that A is a $m \times n$ matrix and consider a linear equation $Ax = b$ and

$$x^* = A^+ b$$

  - If $m \leq n$, then x^* has the smallest Euclidean norm $||x||_2$ among all possible solutions
  - If $m > n$ then $Ax = b$ may not have a solution. In this case x^* minimizes $||Ax - b||_2$

The trace of a matrix

- The trace operator is defined as the sum of all the diagonal elements of a matrix:

$$\mathbf{tr}(A) = \sum_i a_{i,i}$$

- The trace is related to the Frobenius norm of a matrix:

$$||A||_F = \sqrt{\mathbf{tr}(AA^T)}$$

Low-rank matrix approximation

- Given is a matrix $M \in \mathbf{R}^{m \times n}$ and a positive integer r
- The matrix approximation problem:

$$\begin{array}{ll} \text{minimize} & \left\| M - \widehat{M} \right\|_F \\ \text{subject to} & \mathbf{rank}\big(\widehat{M}\big) \leq r \end{array}$$

- Let $M = U \Sigma V^T$ be the SVD
- Let Σ_1 be a $r \times r$ diagonal matrix, U_1 be a $m \times r$ matrix and V_1 be a $n \times r$ matrix

$$U = (U_1, U_2), \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \text{ and } V = (V_1, V_2)$$

Matrix approximation lemma

- Theorem: The rank-r matrix $U_1 \Sigma_1 V_1^T$ is a solution to the matrix approximation problem.

  Moreover, it holds:

$$\left\| M - U_1 \Sigma_1 V_1^T \right\|_F = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_n^2}$$

  The solution is unique if $\sigma_r \neq \sigma_{r+1}$

- Comment: if M is a square matrix, then all the above statements hold by using the eigen-decomposition