Notes on Matrix Theory and Basic Linear Algebra

Definitions

- For positive whole numbers m, n, an $m \times n$ real matrix (in $\mathbb{R}^{m\times n}$) A is a rectangular array of real numbers with m rows and n columns. If $m = n$ the matrix is *square.* For example

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ -1 & 5 \end{pmatrix} B = \begin{pmatrix} 1 & 0 & -3 \\ -2 & 1 & 1 \\ 3 & -4 & 0 \end{pmatrix}.$$

  A is 3×2 while B is 3×3.

- The transpose, A^T, of $A \in \mathbb{R}^{m\times n}$ is the matrix in $\mathbb{R}^{n\times m}$ formed by interchanging the rows and columns of A. In our above example

$$A^T = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & 5 \end{pmatrix} B^T = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -4 \\ -3 & 1 & 0 \end{pmatrix}.$$

- The element in the i, j position of a matrix A is written as a_{ij}.

- The sum of two matrices $A, B \in \mathbb{R}^{m\times n}$ is the matrix $C = A + B$ in $\mathbb{R}^{m\times n}$ whose i, j elements are $c_{ij} = a_{ij} + b_{ij}$ for $1 \le i \le m$, $1 \le j \le m$.

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ -1 & 5 \end{pmatrix} B = \begin{pmatrix} 1 & 0 \\ -2 & 1 \\ 3 & -4 \end{pmatrix} A + B = \begin{pmatrix} 3 & 1 \\ -1 & 1 \\ 2 & 1 \end{pmatrix}$$

- The product $C = AB$ of two matrices $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ is the matrix in $\mathbb{R}^{m\times p}$ whose i, j elements are $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 5 \end{pmatrix} B = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & -4 \\ -3 & 1 & 0 \end{pmatrix} AB = \begin{pmatrix} 7 & -2 & 2 \\ -10 & 6 & -5 \end{pmatrix}$$

- I_n denotes the $n \times n$ identity matrix with ones along its diagonal and zeros elsewhere. When the dimension is clear from context, we shall drop the subscript. For all $A \in \mathbb{R}^{n\times n}$, $AI_n = I_n A = A$.

- As a special case, the product of a matrix A in $\mathbb{R}^{m\times n}$ and a vector $v \in \mathbb{R}^n$ is the vector $w = Av$ in $\mathbb{R}^m$ whose i^{th} component is $\sum_{j=1}^{n} a_{ij} v_j$.

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 5 \end{pmatrix} v = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} Av = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Linear independence and rank

- A set of vector $v_1, \ldots, v_p$ in $\mathbb{R}^n$ is linearly dependent if there are real numbers $c_1, \ldots, c_p$, not all zero, such that
$$c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = 0.$$
If a set is not linearly dependent then it is linearly independent.
$$\begin{aligned} \textit{Linearly dependent} : v_1 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} v_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ v_1 &+ v_2 - v_3 = 0. \end{aligned}$$
$$\textit{Linearly independent} : w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} w_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

- The vectors $v_1, \ldots v_p$ is said to span $\mathbb{R}^n$ if every vector in $\mathbb{R}^n$ can be written as a linear combination of $v_1, \ldots, v_p$. In the above example w_1, w_2, w_3 span $\mathbb{R}^3$.

- A basis for $\mathbb{R}^n$ is a set of linearly independent vectors that spans $\mathbb{R}^n$. It is a fundamental fact of linear algebra that all bases contain the same number of elements (n in the case of $\mathbb{R}^n$). This number is the dimension of the space. Above, w_1, w_2, w_3 form a basic for $\mathbb{R}^3$.

- The column rank of a matrix $A \in \mathbb{R}^{m \times n}$, $cr(A)$, is the size of the largest set of linearly independent columns that can be chosen from the matrix. The row rank $rr(A)$ is defined in the same way.

- Obviously $cr(A) \leq n$ and $rr(A) \leq m$. In fact, the row and column rank of a matrix are always equal so $rank(A) = cr(A) = rr(A) \leq \min(m, n)$.

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 5 \end{pmatrix} rank(A) = 2. B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} rank(B) = 3.$$

Determinants and Traces

- The trace of a matrix $A \in \mathbb{R}^{n \times n}$ is the sum of its diagonal elements.
$$A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & -4 \\ -3 & 1 & 0 \end{pmatrix}, trace(A) = 2$$

- The determinant of a 2×2 matrix
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
is $\det(A) = a_{11}a_{22} - a_{21}a_{12}$.

- For an $n \times n$ matrix A, A_{ij} denotes the matrix obtained by removing the i^{th} row and j^{th} column of A. (A_{ij} is $n-1 \times n-1$).
- For $A \in \mathbb{R}^{n \times n}$ $\det(A)$ is given by

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) \cdots + (-1)^{n+1}a_{1n}\det(A_{1n}).$$

$$\begin{aligned} A &= \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & -4 \\ -3 & 1 & 0 \end{pmatrix} \\ \det(A) &= 1det\begin{pmatrix} 1 & -4 \\ 1 & 0 \end{pmatrix} - (-1)det\begin{pmatrix} 2 & -4 \\ -3 & 0 \end{pmatrix} + 3det\begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix} = 7 \end{aligned}$$

Inverses

- A in $\mathbb{R}^{n \times n}$ is invertible if there is a matrix A^{-1} in $\mathbb{R}^{n \times n}$ satisfying $AA^{-1} = A^{-1}A = I$.
- A is invertible if and only if $rank(A) = n$.
- A is invertible if and only if $\det(A) \neq 0$.
- A is invertible if and only if the only vector satisfying $Av = 0$ is $v = 0$.
- A key fact is that for $A, B \in \mathbb{R}^{n \times n}$, $\det(AB) = \det(A)\det(B)$.

Eigenvalues and Eigenvectors

- If $Ax = \lambda x$ for some non-zero vector $x \in \mathbb{C}^n$ and some $\lambda \in \mathbb{C}$, then x is an eigenvector of A corresponding to the eigenvalue λ.

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

  Eigenvalue $= 3$, Eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

- If A is real then for every real eigenvalue there is at least one real eigenvector.
- λ is an eigenvalue of A if and only if it satisfies $\det(A - \lambda I) = 0$.
- $A \in \mathbb{R}^{n \times n}$ has n eigenvalues in $\mathbb{C}$, which need not all be distinct. The number of times λ occurs as a root of $\det(A - tI) = 0$ is the *algebraic multiplicity* of λ. The set of all eigenvalues of A is the spectrum of A and is denoted $\sigma(A)$.
- The space of all vectors $v \in \mathbb{C}^n$ satisfying $Av = \lambda v$ is the λ-eigenspace of A. The dimension of this space is the *geometric multiplicity* of λ.
- If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A \in \mathbb{R}^{n \times n}$ (counting multiplicities so some may be repeated) then $\det(A) = \lambda_1\lambda_2 \cdots \lambda_n$ and $\mathrm{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

- If $B = SAS^{-1}$ for an invertible $S \in \mathbb{R}^{n\times n}$, then B and A are said to be similar. In this case, $\det(A) = \det(B)$, $trace(A) = trace(B)$ and $\sigma(A) = \sigma(B)$.

Some special classes of matrices

- $D \in \mathbb{R}^{n\times n}$ is diagonal if $d_{ij} = 0$ when $i \neq j$. The eigenvalues of a diagonal matrix are simply the elements along the diagonal.
- $U \in \mathbb{R}^{n\times n}$ is upper triangular if $t_{ij} = 0$ when $i > j$. L is lower triangular if $l_{ij} = 0$ when $i < j$. The eigenvalues of a triangular matrix are given by its diagonal elements.
- A matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if $S^T = S$. All of the eigenvalues of symmetric matrices are real. $A \in \mathbb{C}^{n\times n}$ is Hermitian if $A^* = A$ where A^* is $\bar{A}^T$. As with real symmetric matrices, all eigenvalues of a Hermitian matrix are real.
- $O \in \mathbb{R}^{n\times n}$ is orthogonal if $OO^T = I$.
- If $A \in \mathbb{R}^{n\times n}$ is symmetric then there is an orthogonal matrix $O \in \mathbb{R}^{n\times n}$ and a diagonal matrix $D \in \mathbb{R}^{n\times n}$ such that $A = ODO^T$.