Chapter 1

Linear Algebra

1.1 Example

Consider the following problem: fit a polynomial of degree $n-1$ to the following points:

$$(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$$

where x_i and y_i are real numbers ($x_i, y_i \in \mathbf{R}$). Notice that we are looking for the coefficients of the polynomial

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

and we know that

$$\begin{array}{rcl} p_n(x_1) & = & a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} & = & y_1 \\ p_n(x_2) & = & a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} & = & y_2 \\ & \vdots & & & \\ p_n(x_n) & = & a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} & = & y_n \end{array}$$

Thus, we must solve this linear system of equations for $\{a_0, \ldots, a_{n-1}\}$.

Notice the above set of linear equations can be rewritten as

$$a_0 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + a_1 \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + a_2 \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix} + \cdots + a_{n-1} \begin{pmatrix} x_1^{n-1} \\ x_2^{n-1} \\ \vdots \\ x_n^{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

or, in matrix notation:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

We call the two-dimensional array

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

an $n \times n$ matrix,

$$a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}$$

the vector of unknowns, and

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

the right-hand-side vector.

We are thus left to compute a from the *linear equation*

$$Xa = y$$

1.2 Vectors

A *(column) vector*, x, is the n-tuple of real or complex numbers

$$x = \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix}$$

Here χ_i are called the *components* of vector x. We will denote the set of all vectors with real components $\mathbf{R}^n$ and with complex components $\mathbf{C}^n$.

A *row vector*, x^T, is the n=tuple of real or complex numbers

$$x = \left(\chi_1 \,\middle|\, \chi_2 \,\middle|\, \vdots \,\middle|\, \chi_n \right)$$

Here x^T indicates a transposed (column) vector. (We will always assume vectors are column vectors, unless transposed like this, or explicitly noted.) More about transposition next.

1.3 Matrices

An $m \times n$ matrix, A, is the array

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

with m rows and n columns. The (i,j) component, or *element*, of A refers to α_{ij}, which may be real or complex. The numbers m and n are the dimensions of A. If $m = n$ then the matrix is said to be **square**. Otherwise, it is said to be **rectangular**. We will use the convension of Stewart to use greek letters for real or complex numbers, lowercase italicized letters for vectors, and uppercase italicized letters for matrices.

1.4 Operations on vectors and matrices

1.4.1 Scalar times vector multiplication

$$\alpha x = \alpha \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix} = \begin{pmatrix} \alpha\chi_1 \\ \alpha\chi_2 \\ \vdots \\ \alpha\chi_n \end{pmatrix}$$

1.4.2 Adding vectors

$$x + y = \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} \chi_1 + \eta_1 \\ \chi_2 + \eta_2 \\ \vdots \\ \chi_n + \eta_n \end{pmatrix}$$

1.4.3 Transposition

Definition 1 *Let $A \in \mathbf{R}^{m\times n}$ and $B \in \mathbf{R}^{n\times m}$. Then B is said to be the* transpose *of A if $\beta_{ij} = \alpha_{ji}$ for all $i, i = 1,\ldots,n$ and $j, j = 1,\ldots,m$. Notations:*

$$B = A^T$$

Notice: if $B = A^T$ then $A = B^T$.

1.4.4 Inner product

The *inner product* of two vectors $x, y \in \mathbf{R}^n$ is given by

$$\sum_{k=1}^{n} \chi_i\eta_i = \chi_1\eta_1 + \chi_2\eta_2 + \cdots + \chi_n\eta_n$$

1.4.5 Scalar times a matrix

$$\beta A = \beta \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} = \begin{pmatrix} \beta\alpha_{11} & \beta\alpha_{12} & \cdots & \beta\alpha_{1n} \\ \beta\alpha_{21} & \beta\alpha_{22} & \cdots & \beta\alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \beta\alpha_{m1} & \beta\alpha_{m2} & \cdots & \beta\alpha_{mn} \end{pmatrix}$$

1.4.6 Matrix times a vector

$$\begin{aligned} Ax &= \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix} \\ &= \chi_1 \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{pmatrix} + \chi_2 \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{n2} \end{pmatrix} + \cdots + \chi_n \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{nn} \end{pmatrix} \end{aligned}$$

$$
= \begin{pmatrix} \alpha_{11}\chi_1 + \alpha_{12}\chi_2 + \cdots + \alpha_{1n}\chi_n \\ \alpha_{21}\chi_1 + \alpha_{22}\chi_2 + \cdots + \alpha_{2n}\chi_n \\ \vdots \\ \alpha_{m1}\chi_1 + \alpha_{m2}\chi_2 + \cdots + \alpha_{mn}\chi_n \end{pmatrix}
$$

1.4.7 Matrix times a matrix

$$
\begin{aligned}
AB &= \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2k} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mk} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{kn} \end{pmatrix} \\
&= \left(A \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{k1} \end{pmatrix} \;\middle|\; A \begin{pmatrix} \beta_{12} \\ \beta_{22} \\ \vdots \\ \beta_{k2} \end{pmatrix} \;\middle|\; \cdots \;\middle|\; A \begin{pmatrix} \beta_{1n} \\ \beta_{2n} \\ \vdots \\ \beta_{kn} \end{pmatrix} \right) \\
&= \left(\begin{array}{c|c|c|c} \alpha_{11}\beta_{11} + \cdots + \alpha_{1k}\beta_{k1} & \alpha_{11}\beta_{12} + \cdots + \alpha_{1k}\beta_{k2} & \cdots & \alpha_{11}\beta_{1n} + \cdots + \alpha_{1k}\beta_{kn} \\ \hline \alpha_{21}\beta_{11} + \cdots + \alpha_{2k}\beta_{k1} & \alpha_{21}\beta_{12} + \cdots + \alpha_{2k}\beta_{k2} & \cdots & \alpha_{21}\beta_{1n} + \cdots + \alpha_{2k}\beta_{kn} \\ \hline \vdots & \vdots & & \vdots \\ \hline \alpha_{m1}\beta_{11} + \cdots + \alpha_{mk}\beta_{k1} & \alpha_{m1}\beta_{12} + \cdots + \alpha_{mk}\beta_{k2} & \cdots & \alpha_{m1}\beta_{1n} + \cdots + \alpha_{mk}\beta_{kn} \end{array} \right)
\end{aligned}
$$

Note: $c_{ij} = \sum_{l=1}^{k} \alpha_{il}\beta_{jl}$

Note: A (column) vector of length n can be viewed as a $n \times 1$ matrix. A column vector of length n can be viewed as a $1 \times n$ matrix. The innerproduct of two vectors of length n, x and y, can thus be written as

$$
x^T y = \sum_{i=1}
$$

1.4.8 Blocked matrix-matrix multiplication

Definition 2 *A 2×2 blocked partitioning of a given $m \times n$ matrix A is given by*

$$
A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right)
$$

where

$$
\begin{aligned}
A_{11} &\in \mathbf{R}^{m_1 \times n_1} \\
A_{21} &\in \mathbf{R}^{m_2 \times n_1} \\
A_{12} &\in \mathbf{R}^{m_1 \times n_2} \\
A_{22} &\in \mathbf{R}^{m_2 \times n_2}
\end{aligned}
$$

with $m_1 + m_2 = m$ and $n_1 + n_2 = n$.

Lemma 3 *Let A, B, and C all be partitioned so that*

$$
\begin{aligned}
C &= \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array} \right) \\
A &= \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right) \text{ and} \\
B &= \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right)
\end{aligned}
$$

with

$$
\begin{aligned}
C_{11} &\in \mathbf{R}^{m_1 \times n_1} \\
C_{21} &\in \mathbf{R}^{m_2 \times n_1} \\
C_{12} &\in \mathbf{R}^{m_1 \times n_2} \\
C_{22} &\in \mathbf{R}^{m_2 \times n_2} \\
\\
A_{11} &\in \mathbf{R}^{m_1 \times k_1} \\
A_{21} &\in \mathbf{R}^{m_2 \times k_1} \\
A_{12} &\in \mathbf{R}^{m_1 \times k_2} \\
A_{22} &\in \mathbf{R}^{m_2 \times k_2} \\
\\
C_{11} &\in \mathbf{R}^{k_1 \times n_1} \\
C_{21} &\in \mathbf{R}^{k_2 \times n_1} \\
C_{12} &\in \mathbf{R}^{k_1 \times n_2} \\
C_{22} &\in \mathbf{R}^{k_2 \times n_2}
\end{aligned}
$$

with $m_1 + m_2 = m$, $n_1 + n_2 = n$, and $k_1 + k_2 = k$. Then

$$
C = \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array} \right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right) \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right) = \left(\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array} \right)
$$

Exercise: Check the above result!

Notice: Blocked matrix-matrix multiplication is just like regular matrix-matrix multiplication, except with blocks instead of scalars.

Generalization:

Definition 4 *A $M \times N$ blocked partitioning of a given $m \times n$ matrix A is given by*

$$A = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1N} \\ \hline A_{21} & A_{22} & \cdots & A_{2N} \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{M1} & A_{M2} & \cdots & A_{MN} \end{array} \right)$$

where

$$A_{ij} \quad \in \quad \mathbf{R}^{m_i \times n_j}$$

with $m_1 + m_2 + \cdots + m_M = m$ and $n_1 + n_2 + \cdots + n_N = n$.

Lemma 5 *Let A, B, and C partitioned so that C is blocked into $M \times N$ blocks, A is blocked into $M \times K$ blocks, andn B is blocked into $K \times N$ blocks, with C_{ij} of size $m_i \times n_i$, A_{ij} of size $m_i \times k_j$ and B_{ij} of size $k_i \times n_j$. Then*

$$C_{ij} = \sum_{l=1}^{K} A_{il} B_{lj}$$

Again, blocked matrix-matrix multiplication is just like regular matrix-matrix multiplication, with scalars replaced by matrix blocks.

Notice: Given two matrices A and B, it is generally **NOT** the case that $AB = BA$.

1.4.9 Blocked matrix-vector multiplication

Notice that a vector of length n can be viewed as a $n \times 1$ matrix, and thus we can apply the blocked matrix-matrix multiplication to matrix-vector multiplication:

Corollary 6 *Let A be an $m \times n$ matrix, and x be a vector of length n. Block A into $M \times N$ blocks, $\{A_{ij}\}$ and x into N subvectors $\{x_j\}$. Let A_{ij} be $m_i \times n_j$ and x_j be of length n_j. Then*

$$Ax = \left(\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1N} \\ \hline A_{21} & A_{22} & \cdots & A_{2N} \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{M1} & A_{M2} & \cdots & A_{MN} \end{array} \right) \left(\begin{array}{c} x_1 \\ \hline x_2 \\ \hline \vdots \\ \hline x_n \end{array} \right) = \left(\begin{array}{c} A_{11}x_1 + \cdots + A_{1N}x_N \\ \hline A_{21}x_1 + \cdots + A_{2N}x_N \\ \hline \vdots \\ \hline A_{M1}x_1 + \cdots + A_{MN}x_N \end{array} \right)$$

Often, it will be convenient to partion a $m \times n$ matrix by columns:

$$A = \left(\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_n \end{array} \right)$$

where a_j represents the *vector* that makes up the jth column. Thus

$$\begin{aligned} Ax &= \left(\begin{array}{c|c|c|c} a_1 & a_2 & \vdots & a_n \end{array} \right) \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix} \\ &= \chi_1 a_1 + \chi_2 a_2 + \cdots + \chi_n a_n \end{aligned}$$

1.5 Review of basic linear algebra

Theorem 7 *Let $a, b, c \in \mathbf{R}^n$. Then*

1. $a + b = b + a$ *(Commutativity)*
2. $(a + b) + c = a + (b + c)$ *(Associativity)*

Definition 8 *The* zero vector *is the vector all of whose components equal zero.*

We typically denote a zero vector by "0".

Theorem 9 *Let $a \in \mathbf{R}^n$. Then*

1. $a + 0 = a$
2. *There exists a vector, denoted by $-a$ in $\mathbf{R}^n$ such that $a + (-a) = 0$.*

Definition 10 $a - b = a + (-b)$

Theorem 11 *Let $\lambda, \mu \in \mathbf{R}$ and $a, b \in \mathbf{R}^n$. Then*

1. $(\lambda\mu)a = \lambda(\mu a)$
2. $(\lambda + \mu)a = \lambda a + \mu a$
3. $\lambda(a + b) = \lambda a + \lambda b$
4. $1 \cdot a = a$

1.5.1 Linear independence, subspaces, and bases

Definition 12 *Let* $\{a_i\}_{i=1}^m = \{a_1, a_2, \ldots, a_n\}$ *be a set of vectors of length* m*. Then* y *is said to be a* linear combination *of these vectors is there exist scalars* $\chi_1, \ldots, \chi_n$ *so that* $y = \chi_1 a_1 + \chi_2 a_2 + \cdots + \chi_n a_n$*. The numbers* $\chi_1, \ldots, \chi_n$ *are called the* coefficients*. If* $\chi_1 = \chi_2 = \ldots = \chi_n = 0$ *then the linear combination is said to be* trivial.

Notice that this is equivalent to saying that there exist a vector x of length n such that

$$\begin{aligned} y &= Ax = \left(\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_n \end{array} \right) \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{pmatrix} \\ &= \chi_1 a_1 + \chi_2 a_2 + \cdots + \chi_n a_n \end{aligned}$$

Notice that if $x = 0$, the linear combination is trivial, and $y = 0$.