A Review of Basic Linear Algebra

\[ f_1(x_1, x_2, \ldots, x_n) = 0 \]
\[ f_2(x_1, x_2, \ldots, x_n) = 0 \]
\[ \vdots \]
\[ f_n(x_1, x_2, \ldots, x_n) = 0 \]

\[ x_1^2 + x_2^2 + x_3^2 = 9 \]
\[ x_1 x_2 x_3 = 1 \]
\[ x_1 + x_2 - x_3^2 = 0 \]

\[ x^0 = \left( x_1^0, x_2^0, \ldots, x_n^0 \right)^T \]
\[ f_1(x) \approx f_1(x^0) + \nabla f_1(x^0)(x - x^0) \]
\[ f_2(x) \approx f_2(x^0) + \nabla f_2(x^0)(x - x^0) \]
\[ \vdots \]
\[ f_n(x) \approx f_n(x^0) + \nabla f_n(x^0)(x - x^0) \]
\( f_1(x^0) + \nabla f_1(x^0)(x - x^0) = 0 \)
\( f_2(x^0) + \nabla f_2(x^0)(x - x^0) = 0 \)
\[ \vdots \]
\( f_n(x^0) + \nabla f_n(x^0)(x - x^0) = 0 \)

\[ \nabla f_1(x^0)x = -f_1(x^0) + \nabla f_1(x^0)x^0 \]
\[ \nabla f_2(x^0)x = -f_2(x^0) + \nabla f_2(x^0)x^0 \]
\[ \vdots \]
\[ \nabla f_n(x^0)x = -f_n(x^0) + \nabla f_n(x^0)x^0 \]

**Interpolation and Approximation**

\[
\begin{bmatrix} x & y \\ \hline 1 & 1 \\ 2 & -2 \\ 3 & -4 \\ 4 & 3 \\ 5 & -4 \\ 6 & 1 \\ 7 & -2 \\ 8 & 9 \\ \end{bmatrix}
\]
\[ x_i, y_i \quad i = 1, \ldots, n \quad m \leq n - 1 \]
\[ p(x_1) = a_m x_1^m + \cdots + a_1 x_1 + a_0 = y_1 \]
\[ p(x_2) = a_m x_2^m + \cdots + a_1 x_2 + a_0 = y_2 \]
\[ \vdots \]
\[ p(x_n) = a_m x_n^m + \cdots + a_1 x_n + a_0 = y_n \]

The variables are:

\[ a_m, a_{m-1}, \ldots, a_0 \]

1. If \( m + 1 > n \), there are typically many solutions
2. If \( m + 1 < n \), there may be no solution.
3. If \( m + 1 = n \) (a square system), there is, under nonsingularity, a unique solution.

\[ m + 1 = n \]
\[
\begin{bmatrix}
x_1^{n-1} & x_1^{n-2} & \cdots & x_1 & 1 \\
x_2^{n-1} & x_2^{n-2} & \cdots & x_2 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
x_n^{n-1} & x_n^{n-2} & \cdots & x_n & 1
\end{bmatrix}
\begin{bmatrix}
a_{n-1} \\
a_{n-2} \\
a_{n-3} \\
\vdots \\
a_0
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]
\[ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \]

\[ \vdots \]

\[ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \]

\[ Ax = b \]

If \( A \) has an inverse \( \bar{x} = A^{-1}b \)

Not for computation!!

\( A \) is \( nxn \), \( D \) \( nxn \) is an inverse of \( A \) if

\[ AD = I_{nxn} = DA \]

\[ e^1 = (1,0,\cdots,0)', e^2 = (0,1,0,\cdots,0)', \cdots, e^n = (0,\cdots,0,1)' \]

\[ Ax = e^i, i = 1, \ldots, n \]

\[ [x^1, x^2, \ldots, x^n] \]

\[ A[x^1, x^2, \ldots, x^n] = [Ax^1, \ldots, Ax^n] = [e^1, \ldots, e^n] = I_{nxn} \]
Basic Operations:

\[
x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots, x_n)'
\]

\[
\alpha x = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)'
\]

Inner product

\[
x = (x_1, \ldots, x_n)' \text{ and } y = (y_1, \ldots, y_n)'
\]

\[
x' y = \sum_{i=1}^{n} x_i y_i
\]
Matrix-vector product

\[ Ax = \sum_{j=1}^{n} x_j a^j = \begin{bmatrix} a_1 x \\ \vdots \\ a_n x \end{bmatrix} \]

Given the matrix \( A = (a_{i,j}) \)

Its transpose is \( A^t = (a_{j,i}) \)

\[
\begin{bmatrix} 2 & 3 & 1 \\ 1 & 4 & 2 \end{bmatrix}^t = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 1 & 2 \end{bmatrix}
\]
\[ y = (y_1, \cdots, y_m)' \quad A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad mxn \]

\[ y^t A = \sum_{i=1}^{m} y_i a_i = A^t y \]

\[ A, B \quad A \text{ is } mxn \text{ and } B \text{ is } nXk \]

\[ AB = C = (c_{i,j}) \quad c_{i,j} = a_i b^j = \sum_{k=1}^{n} a_{i,k} b_{k,j} \]

The expression
\[ \alpha_1 a^1 + \alpha_2 a^2 + \cdots + \alpha_k a^k \]

is called a linear combination of the \( k \) vectors
\( a^1, a^2, \cdots, a^k \), where each \( a^i \) is an \( m \)-vector
Linear Independence

The $k$ vectors $a^1, a^2, \cdots, a^k$ in $m$-space are said to be linearly independent if

$$\lambda_1 a^1 + \cdots + \lambda_k a^k = 0 \implies \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

$$A = [a^1, \cdots, a^k]$$

$$A\lambda = 0$$

$$\lambda = (\lambda_1, \cdots, \lambda_k)' = (0, 0, \cdots, 0)'$$

Note: If $a^1, \cdots, a^k$ are linearly independent, then
any subset of these vectors is also linearly independent.
Also, none of these vectors can be the zero vector.

Def: If $a^1, \cdots, a^k$ are not linearly independent, they
are said to be linearly dependent.
Note: Any superset of a set of linearly dependent
vectors is also linearly dependent.
**Fact 1:**

Assume the $k$ vectors in $m$-space, $a^1, a^2, \ldots, a^k$ are linearly independent and assume the $m$-vector $a$ can be written as a linear combination of these $k$ vectors:

$$a = \lambda_1 a^1 + \lambda_2 a^2 + \cdots + \lambda_k a^k$$

Then the weights, or coefficients, $\lambda_1, \lambda_2, \ldots, \lambda_k$ are unique.

**Proof:** Suppose, also, $a = \mu_1 a^1 + \mu_2 a^2 + \cdots + \mu_k a^k$ then, by subtraction,

$$0 = (\lambda_1 - \mu_1) a^1 + (\lambda_2 - \mu_2) a^2 + \cdots + (\lambda_k - \mu_k) a^k$$

$$\Rightarrow 0 = \lambda_i - \mu_i \text{ for all } i = 1, 2, \ldots, k$$
Fact 2

A set of \( m+1 \) vectors in \( m \)-space is linearly dependent

That is

\[ \lambda_1 a^1 + \lambda_2 a^2 + \cdots + \lambda_m a^m + \lambda_{m+1} a^{m+1} = 0 \]

\[ \Rightarrow \text{at least one } \lambda_i \neq 0 \]

or, letting

\[ A = [a^1, a^2, a^3, \cdots, a^m, a^{m+1}], mx(m+1) \]

\[ A\lambda = 0 \]

always has a non-zero solution

\[ \lambda \neq 0 \]
or

\[ a_{11} \lambda_1 + \cdots + a_{1m} \lambda_m + a_{1,m+1} \lambda_{m+1} = 0 \]

\[ \cdots \cdots \cdots \cdots \]

\[ a_{m1} \lambda_1 + \cdots + a_{mm} \lambda_m + a_{m,m+1} \lambda_{m+1} = 0 \]

always has a non-zero solution

\[ m=1: \quad a_{11} \lambda_1 + a_{12} \lambda_2 = 0 \]

Inductive Assumption: Assume true for a \((k-1)\times k\) system

\[ a_{11} \lambda_1 + \cdots + a_{1k} \lambda_k + a_{1,k+1} = 0 \]

\[ \vdots \]

\[ a_{k-1,1} \lambda_1 + \cdots + a_{k-1,k} \lambda_k + a_{k-1,k+1} = 0 \]

\[ \cdots \cdots \cdots \cdots \]

\[ a_{k1} \lambda_1 + \cdots + a_{kk} \lambda_k + a_{k,k+1} = 0 \]

If all the coefficients in the last row are zero, then set \( \lambda_{k+1} = 0 \) and we’re back in the \( k-1 \) equation and \( k \) variable case and, therefore, the inductive assumption implies the existence of a non-zero solution. On the other hand, assume the coefficients in the last row are not all zero. Without loss of generality, we may assume that \( a_{k,k+1} \neq 0 \).
By the inductive assumption, the first \( k-1 \) equations in \( k \) variables has a non-zero solution, \( \lambda_1, \ldots, \lambda_k \). Substitute these \( k \) values into the last equation and define \( \overline{\lambda}_{k+1} = -(\overline{a}_1 \lambda_1 + \cdots + \overline{a}_k \lambda_k) \).

\[
\begin{align*}
\overline{a}_{11} \lambda_1 + \cdots + \overline{a}_{1k} \lambda_k + 0 \lambda_{k+1} &= 0 \\
\overline{a}_{k-1,1} \lambda_1 + \cdots + \overline{a}_{k-1,k} \lambda_k + 0 \lambda_{k+1} &= 0 \\
\overline{a}_{kk} \lambda_k + 1 \cdot \lambda_{k+1} &= 0
\end{align*}
\]

**Corollary:** If the set of vectors \( \{a^1, \ldots, a^k\} \) in \( \mathbb{R}^m \) is linearly independent, then \( k \leq m \).

**Replacement**

**Observation:** Let \( \{a^1, \ldots, a^k\} \) be a set of linearly independent vectors. Let \( a \) be a vector that is a linear combination of \( \{a^1, \ldots, a^k\} \). That is

\[
a = \lambda_1 a^1 + \lambda_2 a^2 + \cdots + \lambda_k a^k
\]

Assume \( \lambda \neq 0 \). In particular, assume \( \lambda_i \neq 0 \). Then, the set of vectors

\[
\{a^1, a^2, \ldots a^{-1}, a, a^{i+1}, \ldots, a^k\}
\]

is a linearly independent set. That is, the vector \( a \) may replace the vector \( a^i \) and linear independence is preserved.
Proof: If \{a^1, a^2, \ldots a^{i-1}, a, a^{i+1}, \ldots, a^k\} is a linearly dependent set, there exist numbers \(\mu_1, \mu_2, \ldots, \mu_k\), not all zero, so that
\[
\mu_1 a^1 + \cdots + \mu_{i-1} a^{i-1} + \mu_i a + \mu_{i+1} a^{i+1} + \cdots + \mu_k a^k = 0
\]

\(\mu_i \neq 0\)

\[
a = -\frac{\mu_1}{\mu_i} a^1 - \cdots - \frac{\mu_{i-1}}{\mu_i} a^{i-1} + 0 \cdot a^i - \frac{\mu_{i+1}}{\mu_i} a^{i+1} - \cdots - \frac{\mu_k}{\mu_i} a^k
\]

\(\lambda_i \neq 0\)

Fact 3: Pivot Operation

Let \{a^1, \cdots, a^k\} be vectors in \(\mathbb{R}^m\). Let
\[
a = \lambda_1 a^1 + \lambda_2 a^2 + \cdots + \lambda_i a^i + \cdots + \lambda_k a^k
\]
and let
\[
b = \mu_1 a^1 + \mu_2 a^2 + \cdots + \mu_k a^k
\]
Then, if \(\lambda_i \neq 0\) the vector \(b\) can be written as a linear combination of the vectors \(\{a^1, a^2, \cdots a^{i-1}, a, a^{i+1}, \cdots, a^k\}\).
\[ \lambda_i \neq 0 \]

\[ a^i = \frac{1}{\lambda_i} a - \sum_{j \neq i}^{k} \frac{\lambda_j}{\lambda_i} a^j \]

\[ b = \mu_1 a^1 + \cdots + \mu_i \left[ \frac{1}{\lambda_i} a - \sum_{j \neq i}^{k} \frac{\lambda_j}{\lambda_i} a^j \right] + \cdots + \mu_k a^k \]

\[ = (\mu_1 - \frac{\mu_i}{\lambda_i} \lambda_i) a^1 + (\mu_2 - \frac{\mu_i}{\lambda_i} \lambda_2) a^2 + \cdots + \frac{\mu_i}{\lambda_i} a + \cdots + (\mu_k - \frac{\mu_i}{\lambda_i} \lambda_k) a^k \]

HW:
Assume that \( \lambda_i \neq 0 \); pivot on \( \lambda_i \) and show that the new first column is comprised of precisely the weights derived at the end of Fact 3.
Def: A set of $k$ $m$-vectors $\{a^1, a^2, \ldots, a^k\}$ is said to be a basis for $\mathbb{R}^m$ if

1. The vectors $a^i, i = 1, \ldots, k$, are linearly independent
2. Any vector $a \in \mathbb{R}^m$ is a linearly combination of $a^i, i = 1, \ldots, k$
   (that is $\exists \lambda_1, \ldots, \lambda_k \ni a = \lambda_1 a^1 + \cdots + \lambda_k a^k$)

Ex: $e^1 = (1, 0, 0)^T, e^2 = (0, 1, 0)^T, e^3 = (0, 0, 1)^T$
is a basis for $\mathbb{R}^3$. Then, automatically, $e^1, e^2, e^3$ are linearly independent since

$$
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\lambda_3 \\
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
\end{bmatrix} \Rightarrow \lambda = 0
$$

and

$$x = (x_1, x_2, x_3)^T \iff x = x_1 e^1 + x_2 e^2 + x_3 e^3$$
Similarly, $\mathbb{R}^m$ has a basis; in particular, the vectors

$$e^1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e^2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e^m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

form a basis.

---

**Fact 4:**

Every basis for $\mathbb{R}^m$ contains exactly $m$ vectors.

**Proof:** Assume that $\{a^1, \ldots, a^k\}$ is a set of basis vectors for $\mathbb{R}^m$. Since $\{a^1, \ldots, a^k\}$ is a linearly independent set, we must have $k \leq m$. We assume that $k < m$ and derive a contradiction. Each of the $m$ unit vectors, $e^1, \ldots, e^m$ of $\mathbb{R}^m$ can be written as a linear combination of the vectors in $\{a^1, \ldots, a^k\}$. Then

$$e^1 = \lambda_1 a^1 + \cdots + \lambda_k a^k$$

$\lambda_1 \neq 0$

$$\{e^1, a^2, \ldots, a^k\}$$
\[
e^2 = \mu_1 e^1 + \mu_2 a^2 + \cdots + \mu_k a^k
\]

Without loss of generality
\[
\mu \neq 0, \mu_2 \neq 0 \quad \{e^1, e^2, \cdots, a^k\}
\]

If we continue in this fashion, we'll end up with the statement that each of the vectors \(e^{k+1}, e^{k+2}, \cdots, e^n\) can be written as a linear combination of the vectors \(\{e^1, \cdots, e^k\}\). Of course, the latter is impossible since \(\{e^1, \cdots, e^n\}\) is a linearly independent set. Therefore, we've proven that every basis for \(\mathbb{R}^n\) has exactly \(m\) vectors.

Using virtually the same argument we show that any set of \(m\) linearly independent vectors in \(\mathbb{R}^m\) is a basis.

**Fact 5:**

*A set of \(m\) linearly independent vectors in \(\mathbb{R}^m\) is a basis for \(\mathbb{R}^m\)*

Let \(\{a^1, \cdots, a^m\}\) be a set of linearly independent vectors in \(\mathbb{R}^m\).

Then \(\{a^1, \cdots, a^m\}\) is a basis for \(\mathbb{R}^m\).

**Proof:** We know that \(\{e^1, \cdots, e^m\}\) is a basis for \(\mathbb{R}^m\). Therefore,

\[
a^1 = \lambda_1 e^1 + \cdots + \lambda_m e^m
\]

We know that \(a^1 \neq 0\) since it's a member of a linearly independent set. Therefore, without loss of generality, assume that \(\lambda_i \neq 0\). Then, by the Observation above, the vectors \(\{a^1, e^2, \cdots, e^m\}\) are linearly independent and, by Fact 3, any vector of \(\mathbb{R}^m\) can be written as a linear combination of \(\{a^1, e^2, \cdots, e^m\}\). We continue in this fashion until all of the vectors \(e^1, e^2, \cdots, e^m\) have been replaced by the vectors \(a^1, a^2, \cdots, a^m\).