Linear algebra review

- vector space, subspaces
- independence, basis, dimension
- nullspace and range
- left and right invertibility
a vector space or linear space (over the reals) consists of

- a set \( V \)
- a vector sum \( + : V \times V \rightarrow V \)
- a scalar multiplication \( \cdot : \mathbb{R} \times V \rightarrow V \)
- a distinguished element \( 0 \in V \)

which satisfy a list of properties
Vector space axioms

- $x + y = y + x$, $\forall x, y \in V$  
  + is commutative

- $(x + y) + z = x + (y + z)$, $\forall x, y, z \in V$  
  + is associative

- $0 + x = x$, $\forall x \in V$  
  0 is additive identity

- $\forall x \in V \exists (-x) \in V$ s.t. $x + (-x) = 0$  
  existence of additive inverse

- $(\alpha\beta)x = \alpha(\beta x)$, $\forall \alpha, \beta \in \mathbb{R}$ $\forall x \in V$  
  scalar mult. is associative

- $\alpha(x + y) = \alpha x + \alpha y$, $\forall \alpha \in \mathbb{R}$ $\forall x, y \in V$  
  right distributive rule

- $(\alpha + \beta)x = \alpha x + \beta x$, $\forall \alpha, \beta \in \mathbb{R}$ $\forall x \in V$  
  left distributive rule

- $1x = x$, $\forall x \in V$  
  1 is multiplicative identity
Examples

- $\mathcal{V}_1 = \mathbb{R}^n$, with standard (componentwise) vector addition and scalar multiplication

- $\mathcal{V}_2 = \{0\}$ (where $0 \in \mathbb{R}^n$)

- $\mathcal{V}_3 = \text{span}(v_1, v_2, \ldots, v_k)$ where

  $$\text{span}(v_1, v_2, \ldots, v_k) = \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_i \in \mathbb{R}\}$$

  and $v_1, \ldots, v_k \in \mathbb{R}^n$
Subspaces

- a *subspace* of a vector space is a *subset* of a vector space which is itself a vector space
- roughly speaking, a subspace is closed under vector addition and scalar multiplication
- examples $V_1, V_2, V_3$ above are subspaces of $\mathbb{R}^n$
Vector spaces of functions

\[ V_4 = \{ x : \mathbb{R}_+ \to \mathbb{R}^n \mid x \text{ is differentiable} \} \], where vector sum is sum of functions:

\[(x + z)(t) = x(t) + z(t)\]

and scalar multiplication is defined by

\[(\alpha x)(t) = \alpha x(t)\]

(a point in \( V_4 \) is a trajectory in \( \mathbb{R}^n \))

\[ V_5 = \{ x \in V_4 \mid \dot{x} = Ax \} \]

(points in \( V_5 \) are trajectories of the linear system \( \dot{x} = Ax \))

\[ V_5 \] is a subspace of \( V_4 \)
(Euclidean) norm

for $x \in \mathbb{R}^n$ we define the (Euclidean) norm as

$$
\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}
$$

$\|x\|$ measures length of vector (from origin)

important properties:

- $\|\alpha x\| = |\alpha| \|x\|$ \hspace{1cm} \text{homogeneity}
- $\|x + y\| \leq \|x\| + \|y\|$ \hspace{1cm} \text{triangle inequality}
- $\|x\| \geq 0$ \hspace{1cm} \text{nonnegativity}
- $\|x\| = 0 \iff x = 0$ \hspace{1cm} \text{definiteness}
RMS value and (Euclidean) distance

root-mean-square (RMS) value of vector $x \in \mathbb{R}^n$:

$$\text{rms}(x) = \left( \frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)^{1/2} = \frac{\|x\|}{\sqrt{n}}$$

norm defines distance between vectors: $\text{dist}(x, y) = \|x - y\|$
Independent set of vectors

A set of vectors \( \{v_1, v_2, \ldots, v_k\} \) is *independent* if

\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0 \implies \alpha_1 = \alpha_2 = \cdots = 0
\]

Some equivalent conditions:

- Coefficients of \( \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \) are uniquely determined, *i.e.*,

\[
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k
\]

implies \( \alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_k = \beta_k \)

- No vector \( v_i \) can be expressed as a linear combination of the other vectors \( v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k \)
Basis and dimension

set of vectors \(\{v_1, v_2, \ldots, v_k\}\) is called a \textit{basis} for a vector space \(\mathcal{V}\) if

\[
\mathcal{V} = \text{span}(v_1, v_2, \ldots, v_k)
\]

and

\(\{v_1, v_2, \ldots, v_k\}\) is independent

- equivalently, every \(v \in \mathcal{V}\) \textit{can be uniquely} expressed as

\[
v = \alpha_1 v_1 + \cdots + \alpha_k v_k
\]

- for a given vector space \(\mathcal{V}\), the number of vectors in any basis is the same

- number of vectors in any basis is called the \textit{dimension} of \(\mathcal{V}\), denoted \(\text{dim} \mathcal{V}\)
Nullspace of a matrix

the **nullspace** of $A \in \mathbb{R}^{m \times n}$ is defined as

$$\text{null}(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$$

- **null**$(A)$ is set of vectors mapped to zero by $y = Ax$
- **null**$(A)$ is set of vectors orthogonal to all rows of $A$

**null**$(A)$ gives *ambiguity* in $x$ given $y = Ax$:

- if $y = Ax$ and $z \in \text{null}(A)$, then $y = A(x + z)$
- conversely, if $y = Ax$ and $y = A\tilde{x}$, then $\tilde{x} = x + z$ for some $z \in \text{null}(A)$

**null**$(A)$ is also written $\mathcal{N}(A)$
Zero nullspace

A is called one-to-one if 0 is the only element of its nullspace

\[ \text{null}(A) = \{0\} \]

Equivalently,

- \( x \) can always be uniquely determined from \( y = Ax \) (i.e., the linear transformation \( y = Ax \) doesn’t ‘lose’ information)
- mapping from \( x \) to \( Ax \) is one-to-one: different \( x \)'s map to different \( y \)'s
- columns of \( A \) are independent (hence, a basis for their span)
- \( A \) has a left inverse, i.e., there is a matrix \( B \in \mathbb{R}^{n \times m} \) s.t. \( BA = I \)
- \( A^\top A \) is invertible
Two interpretations of nullspace

suppose $z \in \text{null}(A)$, and $y = Ax$ represents \textit{measurement} of $x$

- $z$ is undetectable from sensors — get zero sensor readings
- $x$ and $x + z$ are indistinguishable from sensors: $Ax = A(x + z)$

$\text{null}(A)$ characterizes \textit{ambiguity} in $x$ from measurement $y = Ax$

alternatively, if $y = Ax$ represents \textit{output} resulting from input $x$

- $z$ is an input with no result
- $x$ and $x + z$ have same result

$\text{null}(A)$ characterizes \textit{freedom of input choice} for given result
Left invertibility and estimation

- apply left-inverse $B$ at output of $A$
- then estimate $\hat{x} = BAx = x$ as desired
- non-unique: both $B$ and $C$ are left inverses of $A$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix}$$
The **range** of a matrix $A \in \mathbb{R}^{m \times n}$ is defined as

\[
\text{range}(A) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m
\]

**range**$(A)$ can be interpreted as

- the set of vectors that can be ‘hit’ by linear mapping $y = Ax$
- the span of columns of $A$
- the set of vectors $y$ for which $Ax = y$ has a solution

**range**$(A)$ is also written $\mathcal{R}(A)$
Onto matrices

A is called \textit{onto} if \textbf{range}(A) = \mathbb{R}^m

equivalently,

\begin{itemize}
  \item $Ax = y$ can be solved in $x$ for any $y$
  \item columns of $A$ span $\mathbb{R}^m$
  \item $A$ has a \textit{right inverse}, i.e., there is a matrix $B \in \mathbb{R}^{n \times m}$ s.t. $AB = I$
  \item rows of $A$ are independent
  \item $\text{null}(A^T) = \{0\}$
  \item $AA^T$ is invertible
\end{itemize}
Interpretations of range

suppose \( v \in \text{range}(A), w \not\in \text{range}(A) \)

\( y = Ax \) represents measurement of \( x \)

- \( y = v \) is a possible or consistent sensor signal
- \( y = w \) is impossible or inconsistent; sensors have failed or model is wrong

\( y = Ax \) represents output resulting from input \( x \)

- \( v \) is a possible result or output
- \( w \) cannot be a result or output

\( \text{range}(A) \) characterizes the possible results or achievable outputs
Right invertibility and control

- apply right-inverse $C$ at *input* of $A$
- then output $y = ACy_{des} = y_{des}$ as desired
Inverse

A ∈ ℝ^{n×n} is invertible or nonsingular if it has both a left and right inverse equivalent conditions:

- columns of A are a basis for ℝ^n
- rows of A are a basis for ℝ^n
- y = Ax has a unique solution x for every y ∈ ℝ^n
- null(A) = {0}
- range(A) = ℝ^n
Inverse

if a matrix $A$ has both a left inverse and a right inverse, then they are equal

\[ BA = I \text{ and } AC = I \implies B = C \]

▶ hence if $A$ is invertible then the inverse is unique

▶ $AA^{-1} = A^{-1}A = I$
Interpretations of inverse

suppose \( A \in \mathbb{R}^{n \times n} \) has inverse \( B = A^{-1} \)

- mapping associated with \( B \) undoes mapping associated with \( A \) (applied either before or after!)

- \( x = By \) is a perfect (pre- or post-) equalizer for the channel \( y = Ax \)

- \( x = By \) is unique solution of \( Ax = y \)
Dual basis interpretation

- Let $a_i$ be columns of $A$, and $\tilde{b}_i^T$ be rows of $B = A^{-1}$

- From $y = x_1a_1 + \cdots + x_na_n$ and $x_i = \tilde{b}_i^T y$, we get

$$y = \sum_{i=1}^{n} (\tilde{b}_i^T y) a_i$$

Thus, inner product with rows of inverse matrix gives the coefficients in the expansion of a vector in the columns of the matrix.

- $\{\tilde{b}_1, \ldots, \tilde{b}_n\}$ and $\{a_1, \ldots, a_n\}$ are called dual bases
Change of coordinates

► standard basis vectors in $\mathbb{R}^n$: $(e_1, e_2, \ldots, e_n)$ where $e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ (1 in $i$th component)

► obviously we have

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

$x_i$ are called the coordinates of $x$ (in the standard basis)

► if $(t_1, t_2, \ldots, t_n)$ is another basis for $\mathbb{R}^n$, we have

$$x = \tilde{x}_1 t_1 + \tilde{x}_2 t_2 + \cdots + \tilde{x}_n t_n$$

where $\tilde{x}_i$ are the coordinates of $x$ in the basis $(t_1, t_2, \ldots, t_n)$

► then $x = T\tilde{x}$ and $\tilde{x} = T^{-1}x$
Similarity transformation

consider linear transformation $y = Ax$, $A \in \mathbb{R}^{n \times n}$
express $y$ and $x$ in terms of $t_1, t_2 \ldots, t_n$, so $x = T\tilde{x}$ and $y = T\tilde{y}$, then

\[
\tilde{y} = (T^{-1}AT)\tilde{x}
\]

- $A \longrightarrow T^{-1}AT$ is called similarity transformation

- similarity transformation by $T$ expresses linear transformation $y = Ax$ in coordinates $t_1, t_2, \ldots, t_n$