I. Preliminaries

A. Definition: a vector space (linear space) consists of:
   1. a field $F$ of scalars. (We are interested in $F = \mathbb{R}$).
   2. a set $V$ of vectors.
   3. an operation $+$, called vector addition, which for all $x, y, z \in V$ satisfies:
      - $x + y \in V$
      - $x + y = y + x$
      - $x + (y + z) = (x + y) + z$
      - $\exists$ a zero vector 0, such that $x + 0 = x$
      - $\exists$ an inverse $-x$, such that $x + (-x) = 0$
   4. an operation $\cdot$, called scalar multiplication, which for all $x, y \in V$ and $\alpha, \beta \in F$ satisfies:
      - $\alpha x \in V$
      - $\exists$ and identity 1, such that $1x = x$
      - $(\alpha \beta)x = \alpha(\beta x)$
      - $\alpha(x + y) = \alpha x + \alpha y$
      - $(\alpha + \beta)x = \alpha x + \beta x$

B. Examples
   1. $V = \mathbb{R}^n = \{x \mid x = (x_1, \ldots, x_n), x_i \in \mathbb{R}\}$
      $x + y = (x_1 + y_1, \ldots, x_n + y_n)$
      $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$
   2. $V = \mathbb{R}^{m \times n} = \{m \times n$ matrices$\}$
      $(A + B)_{ij} = a_{ij} + b_{ij}$
      $(\alpha A)_{ij} = \alpha a_{ij}$
   3. $V = P_n = \{\text{polynomials of degree } \leq n\}$
   4. $V = \{\text{continuous functions on } [0, 1]\}$

C. Definition: a linear combination of the vectors $x_1, \ldots, x_n$ is given by
   $$\sum_{i=1}^{n} \alpha_i x_i = \alpha_1 x_1 + \cdots + \alpha_n x_n,$$
   where $\alpha_1, \ldots, \alpha_n$ are scalars.

D. Definition: a set of vectors $\{x_1, \ldots, x_n\}$ is linearly independent iff
   $$\sum_{i=1}^{n} \alpha_i x_i = 0 \implies \alpha_i = 0$$

E. Basis
   - Definition: a linearly independent set of vectors $\{x_1, \ldots, x_n\}$ is a basis for a vector space $V$ iff for every $x \in V$, there exist scalars $\alpha_1, \ldots, \alpha_n$, such that $x = \sum_{i=1}^{n} \alpha_i x_i$
   - Fact: all bases of a space have the same number of vectors (the dimension of the space).
   - Fact: every vector $x$ has a unique representation in a given basis. Thus we can represent a vector by its coefficients, $x = (\alpha_1, \ldots, \alpha_n)$.
   - terminology: we say that a basis spans a vector space.

F. Definition: given vector spaces $V$ and $W$, a linear transformation is a mapping $A : V \rightarrow W$, such that for all $\alpha, \beta \in \mathbb{R}$ and for all $x, y \in V$,
   $$A(\alpha x + \beta y) = \alpha Ax + \beta Ay.$$
II. Norms

A. Definition: a norm on a vector space $V$ is a function $\| \cdot \| : V \to \mathbb{R}$, such that for all $x, y \in V$ and for all scalars $\alpha$

\[
\begin{align*}
\|x\| & \geq 0 \\
\|x\| &= 0 \text{ iff } x = 0 \\
\|\alpha x\| &= |\alpha| \|x\| \\
\|x + y\| & \leq \|x\| + \|y\| \\
\end{align*}
\]

B. Examples

1. on $\mathbb{R}^n$

\[
\begin{align*}
\|x\|_2 &= (x_1^2 + \cdots + x_n^2)^{1/2} \quad (l_2, \text{Euclidean}) \\
\|x\|_1 &= |x_1| + \cdots + |x_n| \quad (l_1) \\
\|x\|_\infty &= \max_i |x_i| \quad (l_\infty, \text{infinity}) \\
\end{align*}
\]

2. on $\mathbb{R}^{n \times n}$

- defined in terms of a norm on $\mathbb{R}^n$.

\[
\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max \|Ax\|
\]

- intuition: maximum amount that a unit-vector is stretched by the linear transformation represented by $A$.

- examples:

\[
\begin{align*}
\|A\|_2 &= \sqrt{\rho(A^T A)}, \quad \text{where } \rho(A^T A) = \text{max eigenvalue of } A^T A \\
\|A\|_1 &= \max_j \sum_{i=1}^n |A_{ij}| \\
\|A\|_\infty &= \max_i \sum_{j=1}^n |A_{ij}| \\
\end{align*}
\]

3. on $V = \{\text{all polynomials}\}$

\[
\|p\|_\infty = \max_{0 \leq x \leq 1} |p(x)|
\]

III. Inner products

A. Definition: an inner product is a function $(\cdot, \cdot) : V \times V \to \mathbb{R}$, satisfying:

\[
\begin{align*}
(x, x) & \geq 0 \\
(x, x) &= 0 \text{ iff } x = 0 \\
(x, y) &= (y, x) \\
(\alpha x, y) &= \alpha (x, y) \\
(x + y, z) &= (x, z) + (y, z) \\
\end{align*}
\]

B. Examples

\[
\begin{align*}
V &= \mathbb{R}^n & (x, y) &= x^T y = \sum_{i=1}^n x_i y_i \\
V &= P_n = \{\text{polynomials of degree } \leq n\} & (f, g) &= \sum_{i=1}^{n+1} f(x_i) g(x_i) \\
V &= \{\text{continuous functions on } [0, 1]\} & (f, g) &= \int_0^1 f(x) g(x) dx \\
\end{align*}
\]

C. Orthogonality.

- two vectors $x, y$ are orthogonal if $(x, y) = 0$.
- a set $\{x_1, \ldots, x_n\}$ is orthogonal if $(x_i, x_j) = 0, i \neq j$.
- an orthogonal set $\{x_1, \ldots, x_n\}$ is orthonormal if $(x_i, x_i) = 1$.

D. Norms derived from inner products.

\[
\|x\| = \sqrt{(x, x)} \text{ is a norm.}
\]

E. Cauchy-Schwarz inequality:

\[
| (x, y) | \leq \|x\| \|y\|, \quad \text{where } \| \cdot \| = \sqrt{(\cdot, \cdot)}
\]