Part 1a: Inner product, Orthogonality, Vector/Matrix norm

September 19, 2018

1. **Inner product** on a linear space $\mathbb{V}$ over the number field $\mathbb{F}$

- A map $\langle\cdot,\cdot\rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{F}$ is called an *inner product*, if it satisfies the following three conditions for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}$ and for all $\alpha \in \mathbb{F}$:

  (1) Conjugate symmetry:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$$

  (2) Positive definiteness:

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

  (3) Linearity in the first argument:

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle, \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$$

  Example: the standard inner product on the space $\mathbb{V} = \mathbb{C}^m$:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^m, \quad \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i=1}^{m} x_i \bar{y}_i.$$

2. Orthogonality

- Orthogonality is a mathematical concept with respect to a given inner product $\langle \cdot, \cdot \rangle$, i.e., $\mathbf{x}$ and $\mathbf{y}$ are orthogonal means $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

  Example: For any hermitian positive definite matrix $\mathbf{A}$, $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = \mathbf{y}^*\mathbf{A}\mathbf{x}$ is an inner product. If $\mathbf{y}^*\mathbf{A}\mathbf{x} = 0$, we call $\mathbf{x}$ and $\mathbf{y}$ are orthogonal with respect to $\langle \cdot, \cdot \rangle_{\mathbf{A}}$.

3. Orthogonality with respect to the standard inner product

- Two vectors $\mathbf{x}$ and $\mathbf{y}$ are called *orthogonal*: if $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^*\mathbf{x} = 0$
- Two sets of vectors $\mathcal{X}$ and $\mathcal{Y}$ are called orthogonal: if $\forall \mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in \mathcal{Y}$, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^*\mathbf{x} = 0$
- A set of nonzero vectors $\mathcal{S}$ is orthogonal if $\forall \mathbf{x}, \mathbf{y} \in \mathcal{S}$ and $\mathbf{x} \neq \mathbf{y}$, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^*\mathbf{x} = 0$; if further $\forall \mathbf{x} \in \mathcal{S}, \mathbf{x}^*\mathbf{x} = 1$, $\mathcal{S}$ is *orthonormal*.

Proposition 1

The vectors in an orthogonal set $\mathcal{S}$ are linearly independent.

4. Orthogonal components of a vector

- Inner products can be used to decompose arbitrary vectors into orthogonal components. Given *orthonormal* set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ and an arbitrary vector $\mathbf{v}$, let

$$\mathbf{r} = \mathbf{v} - \langle \mathbf{v}, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{v}, \mathbf{q}_2 \rangle \mathbf{q}_2 - \cdots - \langle \mathbf{v}, \mathbf{q}_n \rangle \mathbf{q}_n.$$

  Obviously,

$$\mathbf{r} \in \operatorname{span}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}^{\perp}.$$

  Thus we see that $\mathbf{v}$ can be decomposed into $n+1$ orthogonal components:

$$\mathbf{v} = \mathbf{r} + \langle \mathbf{v}, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{v}, \mathbf{q}_2 \rangle \mathbf{q}_2 + \cdots + \langle \mathbf{v}, \mathbf{q}_n \rangle \mathbf{q}_n.$$

  We call $\langle \mathbf{v}, \mathbf{q}_i \rangle \mathbf{q}_i$ the part of $\mathbf{v}$ in the direction of $\mathbf{q}_i$, and $\mathbf{r}$ the part of $\mathbf{v}$ orthogonal to the subspace $\operatorname{span}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.

Exercise: Write the expression for $\mathbf{v}$ when $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ are only orthogonal.

5. Cauchy-Schwarz inequality

- For any given inner product $\langle \cdot, \cdot \rangle$,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

The equality holds if and only if $\mathbf{x} = \alpha \mathbf{y}$ i,e., $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Exercise: Prove the inequality. Hint: write

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}\mathbf{y} + \mathbf{z}.$$

Then $\langle \mathbf{z}, \mathbf{y} \rangle = 0$. Consider $\langle \mathbf{x}, \mathbf{x} \rangle$.

Application: For any hermitian positive definite matrix $\mathbf{A}$,

$$|\mathbf{y}^* \mathbf{A} \mathbf{x}|^2 \leq (\mathbf{x}^* \mathbf{A} \mathbf{x})(\mathbf{y}^* \mathbf{A} \mathbf{y}).$$

6. Norm on a linear space $\mathbb{V}$ over the number field $\mathbb{F}$ ($\mathbb{C}$ or $\mathbb{R}$)

- A function $\|\cdot\| : \mathbb{V} \to \mathbb{R}$ is called a *norm* if it satisfies the following three conditions for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ and for all $\alpha \in \mathbb{F}$:

  (1) Positive definiteness:

  $$\|\mathbf{x}\| \geq 0, \quad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

  (2) Positive homogeneity:

  $$\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$$

  (3) Triangle inequality:

  $$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Exercise: for any given inner product $\langle \cdot, \cdot \rangle$, let $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$

(1) Prove that the function $\| \cdot \|$ is a norm.

(2) Prove the parallelogram law

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

(3) For a set of n orthogonal (with respect to the inner product $\langle \cdot, \cdot \rangle$) vectors $\{\mathbf{x}_i\}$, prove that

$$\left\| \sum_{i=1}^{n} \mathbf{x}_i \right\|^2 = \sum_{i=1}^{n} \|\mathbf{x}_i\|^2.$$

- The function $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ is called the norm *indued* by the inner product $\langle \cdot, \cdot \rangle$. Correspondingly, the Cauchy-Schwarz inequality becomes

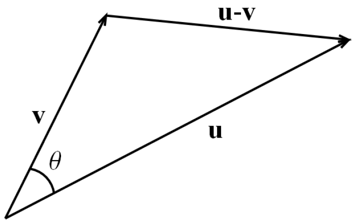
$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

- **Geometric interpretation of an inner product**

  For simplicity, we will consider vectors in $\mathbb{R}^2$. Let $\|\cdot\|$ be the norm induced by the inner product $\langle\cdot,\cdot\rangle$, i.e.,

$$\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}, \quad \text{then} \quad \langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta,$$

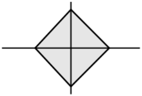
  where θ is the angle between vectors $\mathbf{u}$ and $\mathbf{v}$ such that $0 \le \theta \le \pi$.



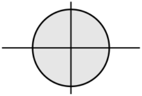
7. **Vector norm** on $\mathbb{C}^m$

- p-norm

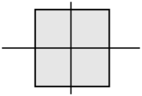
$$\|x\|_1 = \sum_{i=1}^{m} |x_i|,$$



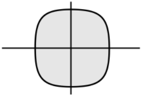
$$\|x\|_2 = \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2} = \sqrt{x^* x},$$



$$\|x\|_\infty = \max_{1 \le i \le m} |x_i|,$$



$$\|x\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p} \qquad (1 \le p < \infty).$$



  For $p = 2$, the norm $\| \cdot \|_2$ on $\mathbb{C}^m$ is also called the Euclidean norm.

- Matlab: `norm` for $1, 2, \infty$ norms

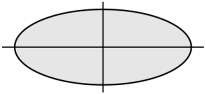
- Weighted norm

  Let $\|\cdot\|$ denote any norm on $\mathbb{C}^m$. Suppose a diagonal matrix $\mathbf{W} = \mathrm{diag}\{w_1, \cdots, w_m\}$, $w_i \neq 0$. Then

$$\|\mathbf{x}\|_{\mathbf{W}} = \|\mathbf{W}\mathbf{x}\|$$

  is a norm, called *weighted norm*. For example, weighted 2-norm

$$\|x\|_W = \left(\sum_{i=1}^{m} |w_i x_i|^2 \right)^{1/2}.$$



- Dual norm

  Let $\|\cdot\|$ denote any norm on $\mathbb{C}^m$. The corresponding *dual norm* $\|\cdot\|'$ is defined by the formula

$$\|\mathbf{x}\|' = \sup_{\|\mathbf{y}\|=1} |\mathbf{y}^*\mathbf{x}|.$$

8. **Matrix norm** on $\mathbb{C}^{m\times n}$

- Frobenius norm: $\forall \mathbf{A} \in \mathbb{C}^{m\times n}$, define

$$\|\mathbf{A}\|_{\mathrm{F}} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2} = \left(\sum_{j=1}^{n} \|\mathbf{a}_j\|_2^2\right)^{1/2}$$

or

$$\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(\mathbf{A}^*\mathbf{A})} = \sqrt{\mathrm{tr}(\mathbf{A}\mathbf{A}^*)}.$$

- Max norm:

$$\|\mathbf{A}\|_{\max} = \max_{i,j} |a_{ij}|.$$

- Induced matrix norm (operator norm): $\forall \mathbf{A} \in \mathbb{C}^{m\times n}$, define

$$\|\mathbf{A}\|_{\alpha,\beta} = \sup_{\substack{\mathbf{x}\in\mathbb{C}^n \\ \mathbf{x}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\beta} = \sup_{\substack{\mathbf{x}\in\mathbb{C}^n \\ \|\mathbf{x}\|_\beta=1}} \|\mathbf{A}\mathbf{x}\|_\alpha = \sup_{\substack{\mathbf{x}\in\mathbb{C}^n \\ \|\mathbf{x}\|_\beta\leq 1}} \|\mathbf{A}\mathbf{x}\|_\alpha,$$

where $\|\cdot\|_\alpha$ is a norm on $\mathbb{C}^m$ and $\|\cdot\|_\beta$ is a norm on $\mathbb{C}^n$. We say that $\|\cdot\|_{\alpha,\beta}$ is the matrix norm induced by $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$.

Exercise: $\forall \mathbf{x} \in \mathbb{C}^n$, prove that $\|\mathbf{A}\mathbf{x}\|_\alpha \le \|\mathbf{A}\|_{\alpha,\beta}\|\mathbf{x}\|_\beta$.

Exercise: Let $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times r}$ and let $\|\cdot\|_\alpha$, $\|\cdot\|_\beta$, and $\|\cdot\|_\gamma$ be norms on $\mathbb{C}^m, \mathbb{C}^n$, and $\mathbb{C}^r$, rspectively. Prove the induced matrix norms $\|\cdot\|_{\alpha,\gamma}$, $\|\cdot\|_{\alpha,\beta}$, and $\|\cdot\|_{\beta,\gamma}$ satisfy

$$\|\mathbf{A}\mathbf{B}\|_{\alpha,\gamma} \le \|\mathbf{A}\|_{\alpha,\beta}\|\mathbf{B}\|_{\beta,\gamma}.$$

Exercise: Prove that $\|\mathbf{A}\|_{\infty,1} = \max_{i,j}|a_{ij}| = \|\mathbf{A}\|_{\max}$

- The Frobenius norm $\|\cdot\|_{\mathrm{F}}$ on $\mathbb{C}^{m\times n}$ is not induced by norms on $\mathbb{C}^m$ and $\mathbb{C}^n$. See the following references,

  [1]. Vijaya-Sekhar Chellaboina and Wassim M. Haddad
  Is the Frobenius matrix norm induced?
  IEEE Trans. Automat. Control 40 (1995), no. 12, 2137–2139.

  [2]. Djouadi, Seddik M.
  Comment on: "Is the Frobenius matrix norm induced?" With a reply by Chellaboina and Haddad.

  IEEE Trans. Automat. Control 48 (2003), no. 3, 518–520.

9. Induced matrix p-norm of $\mathbf{A} \in \mathbb{C}^{m \times n}$

- For $p \in [1, +\infty]$,

$$\|\mathbf{A}\|_p := \|\mathbf{A}\|_{p,p} = \sup_{\mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_p = 1} \|\mathbf{A}\mathbf{x}\|_p.$$

Example: p-norm of a diagonal matrix $\mathbf{D} = \mathrm{diag}\{d_1, \cdots, d_m\}$

$$\|\mathbf{D}\|_p = \max_{1 \le i \le m} |d_i|$$

Example: $1, 2, \infty$-norm

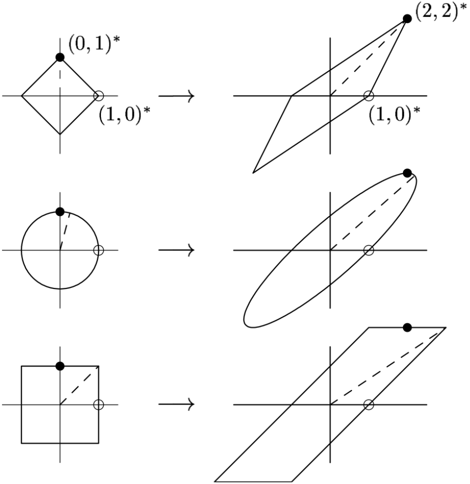
$$\|\mathbf{A}\|_1 = \max_j \sum_i |a_{ij}|, \quad \|\mathbf{A}\|_\infty = \max_i \sum_j |a_{ij}|$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^*\mathbf{A})} = \sqrt{\lambda_{\max}(\mathbf{A}\mathbf{A}^*)} \le \|\mathbf{A}\|_{\mathrm{F}}$$

For $p = 2$, the norm $\|\cdot\|_2$ on $\mathbb{C}^{m \times n}$ is also called the spectral norm.

- Matlab: `norm` for $1, 2, \infty$-norm

Example: $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$



1-norm: $\longrightarrow$ $\|A\|_1 = 4$

2-norm: $\longrightarrow$ $\|A\|_2 \approx 2.9208$

∞-norm: $\longrightarrow$ $\|A\|_\infty = 3$

10. Unitary invariance $\forall \mathbf{A} \in \mathbb{C}^{m \times n}$

- If $\mathbf{P}$ has orthonormal columns, i.e.,

$$\mathbf{P} \in \mathbb{C}^{p \times m}, \quad p \geq m, \quad \mathbf{P}^*\mathbf{P} = \mathbf{I}_m,$$

  then

$$\|\mathbf{PA}\|_2 = \|\mathbf{A}\|_2, \quad \|\mathbf{PA}\|_{\mathrm{F}} = \|\mathbf{A}\|_{\mathrm{F}}.$$

- If $\mathbf{Q}$ has orthonormal rows, i.e.,

$$\mathbf{Q} \in \mathbb{C}^{n \times q}, \quad n \leq q, \quad \mathbf{Q}\mathbf{Q}^* = \mathbf{I}_n,$$

  then

$$\|\mathbf{AQ}\|_2 = \|\mathbf{A}\|_2, \quad \|\mathbf{AQ}\|_{\mathrm{F}} = \|\mathbf{A}\|_{\mathrm{F}}.$$

11. Unitary matrix

- For $\mathbf{Q} \in \mathbb{C}^{m\times m}$, if $\mathbf{Q}^* = \mathbf{Q}^{-1}$, i.e., $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$, $\mathbf{Q}$ is called *unitary* (or *orthogonal* in the real case).

$$\begin{bmatrix} q_1^* \\ \hline q_2^* \\ \hline \vdots \\ \hline q_m^* \end{bmatrix} \left[\begin{array}{c|c|c|c} q_1 & q_2 & \cdots & q_m \end{array}\right] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

Exercise: Let $\mathbf{Q} \in \mathbb{C}^{m\times m}$ be a unitary matrix. Prove

$$\|\mathbf{Q}\|_2 = 1, \quad \|\mathbf{Q}\|_{\mathrm{F}} = \sqrt{m}.$$