vv214: Orthogonality. Gram-Schmidt Orthogonalization.

Dr.Olga Danilkina

UM-SJTU Joint Institute

June 20, 2019

1. Norms. Normed linear spaces. Convergence.
2. Inner product. Inner product spaces.
3. Natural norm.
4. Banach and Hilbert spaces.
5. The Cauchy-Schwarz inequality.
6. Orthogonal and orthonormal elements of inner product spaces.
7. Orthogonal complements and direct sums.
8. Formal definition of a Fourier series.
9. Correlation.
10. Construction of orthonormal bases. Gram-Schmidt process. QR - factorization.

Norms

$|x| \geq 0$

标量乘

Definition: Let X be a linear space over a scalar field $\mathbb{K}$. A real-valued function $||\cdot||: X \to \mathbb{R}$ defined on X is called a norm provided

1. $||x|| \geq 0 \quad \forall x \in X$ and $||x|| = 0$ iff $x = 0$
2. $||\alpha x|| = |\alpha| ||x|| \quad \forall x \in X \, \forall \alpha \in \mathbb{K}$
3. $||x + y|| \leq ||x|| + ||y|| \quad \forall x, y \in X$

Definition: A linear space with a norm $(X, ||\cdot||)$ is called a normed linear space.

Convergence in a normed linear space is defined as convergence w.r.t norm:

$$\{x_n\} \to x \text{ as } n \to \infty \Rightarrow ||x_n - x|| \to 0 \text{ as } n \to \infty$$

Norms: Examples

$(p,g) = \int_{-1}^{1} p(t)g(t)dt$

1. $\mathbb{R}$: $||x|| = |x|$
2. $\mathbb{R}^n$: $||\bar{x}||_\infty = \max_{i=1..n} |x_i|$, $||\bar{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$, $p \geq 1$
3. $C[a,b]$: $||f(t)|| = \max_{t\in[a,b]} |f(t)|$, $\quad ||f(t)|| = \int_a^b |f(t)|\, dt$
4. l^∞: $||\bar{x}|| = \sup_{i=1..\infty} |x_i|$
5. l^p: $= ||\bar{x}||_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$, $p \geq 1$

Inner Product

$\bar{x} \cdot \bar{y} = (\bar{x}, \bar{y}) = x_1 y_1 + \dots + x_n y_n$

Definition: Let X be a linear space over a scalar field $\mathbb{K}$. A complex-valued function of two variables $(\cdot, \cdot) \colon X \times X \to \mathbb{C}$ is called an inner product (dot product) provided

1. $(x, x) \geq 0 \quad \forall x \in X$ and $(x, x) = 0$ iff $x = 0$
2. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) \quad \forall x, y, z \in X \, \forall \alpha, \beta \in \mathbb{K}$
3. $(x, y) = \overline{(y, x)} \quad \forall x, y \in X$

Examples:

1. $\mathbb{R}^n \colon \; (\bar{x}, \bar{y}) = \sum_{i=1}^{n} x_i y_i, \quad \mathbb{C}^n \colon \; (\bar{x}, \bar{y}) = \sum_{i=1}^{n} x_i \bar{y}_i$
2. $C[a, b] \colon \; (f(t), g(t)) = \int_a^b f(t) g(t) \, dt$
3. $l^2 \colon \; (\bar{x}, \bar{y}) = \sum_{i=1}^{\infty} x_i \bar{y}_i$
4. $M_{n \times n} \colon (A, \, B) = tr(\bar{A}^T B)$ Frobenius inner product

The Cauchy-Schwarz Inequality

Lemma (the Cauchy-Schwarz inequality): Let X be a linear space with an inner product.

$$|(x,y)| \leq \sqrt{(x,x)} \cdot \sqrt{(y,y)}$$

Proof: in class

Lemma (triangle inequality):

$$||x+y|| \leq ||x|| + ||y||$$

Natural Norm

Definition: If X is a linear space with a norm $||\cdot||$, i.e. a normed linear space, then the norm is a natural norm induced by the inner product if

$$||x|| = \sqrt{(x,x)} \quad \forall x \in X$$

Examples: In $\mathbb{R}^n$, $||x|| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$ is the natural norm.

In $C[a,b]$, $||f|| = \left(\int_a^b f^2(t)\,dt\right)^{1/2}$ is the natural norm.

Remark: To verify that a norm $||\cdot||$ is a natural norm, check whether the Parallelogram Equality

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

holds.

In every parallelogram, the sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides.

Important Definitions

Definition: A sequence $\{x_n\} \subset X$ is said to be a Cauchy sequence if

$$||x_n - x_m|| \to 0 \text{ as } n,\ m \to \infty$$

Definition: A normed linear space X is said to be complete if ANY Cauchy sequence converges w.r.t norm in X.

Definition: A complete normed liner space is called a Banach space.

A complete inner product space is called a $\boxed{\text{Hilbert space.}}$

Orthonormality

Definition: Let X be a linear space. Two elements $x, y \in X$ are said to be orthogonal (perpendicular) if $(x, y) = 0$.
Definition: If elements x, y are orthogonal and $||x|| = ||y|| = 1$, then x and y are called orthonormal.
Remark: For any $x \in X, x \neq 0$, the element $y = \dfrac{x}{||x||}$ has the norm $||y|| = 1$. Any element $x \in X$ with $||x|| = 1$ is a unit element.
Examples of orthonormal elements:

1. $\mathbb{R}^n$: $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n$
2. $\mathbb{R}^2$: $\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$
3. $C[-\pi, \pi]$: $\dfrac{1}{\sqrt{2\pi}}, \dfrac{1}{\sqrt{\pi}}\cos t, \dfrac{1}{\sqrt{\pi}}\sin t, \dfrac{1}{\sqrt{\pi}}\cos 2t, \dfrac{1}{\sqrt{\pi}}\sin 2t, \dots$
4. $\mathbb{R}^4$: $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}), (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}),$
$(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$

Orthonormal Basis

Lemma (Pythagorean Theorem): If $x, y \in X$ are orthogonal then

$$||x+y||^2 = ||x||^2 + ||y||^2$$

Lemma (orthogonal decomposition): Let $x, y \in X$, $y \neq 0$.

$$x = cy + w, \quad (w \perp y)$$

$$\text{where } (w, y) = 0 \text{ and } c = \frac{(x,y)}{||y||^2} \Rightarrow w = x - \frac{(x,y)}{||y||^2} y$$

Remark: Prove the Cauchy-Schwarz inequality using an orthogonal decomposition.

Theorem: Orthonormal elements are linearly independent.
Proof: in class

Remark: Any n orthonormal vectors $\bar{u}_1, \ldots, \bar{u}_n \in \mathbb{R}^n$ form a basis of $\mathbb{R}^n$.

Orthonormal Basis

Definition: Let $V \subset X$ be a linear subspace. The set

$$V^{\perp} = \{x \in X : (x, v) = 0 \,\forall v \in V\}$$

is called the orthogonal complement of V.
Is $V^{\perp}$ a linear space as well? Yes.
Definition: A linear space X is said to be a direct sum of its linear subspaces $X_1, X_2 \subset X$ if

$$\forall x \in X \,\exists! \quad x = x_1 + x_2, \, x_1 \in X_1, \, x_2 \in X_2$$

$$X = X_1 \oplus X_2$$

Theorem: If V is a linear subspace of $\mathbb{R}^n$, then

$$\mathbb{R}^n = V \oplus V^{\perp}$$

Proof: in class

Remark: If V is a linear subspace of $\mathbb{R}^n$ with an orthonormal basis $\bar{u}_1, \ldots, \bar{u}_m$, then

$$proj_v \, \bar{x} = (\bar{x}, \bar{u}_1)\bar{u}_1 + \ldots + (\bar{x}, \bar{u}_m)\bar{u}_m$$

Correlation

Def: Let $\bar{x}, \bar{y} \in \mathbb{R}^n$. There is a positive correlation between $\bar{x}$ and $\bar{y}$ if and only if $(\bar{x}, \bar{y}) > 0$.

Def: The correlation coefficient r of two vectors $\bar{x}$ and $\bar{y}$ is

$$r = \cos(\bar{x}, \bar{y}) = \frac{(\bar{x}, \bar{y})}{|\bar{x}||\bar{y}|}$$

Remark: By the Cauchy-Schwarz inequality,

$$|(\bar{x}, \bar{y})| \leq |\bar{x}||\bar{y}| \Rightarrow -1 \leq r \leq 1$$

Correlation: Example

Consider meat consumption and incidence of cancer rate in the following countries:

Country	Consumption	Rate	Deviation: Cons	Deviation: Rate
Japan	26	7.5	-122	-10.7
Finland	101	9.8	-47	-8.4
Israel	124	16.4	-24	-1.8
GB	205	23.3	57	5.1
US	284	34	136	15.8
Mean	148	18.2		

The correlation coefficient is

$$r = \frac{122 \cdot 10.7 + 47 \cdot 8.4 + 24 \cdot 1.8 + 57 \cdot 5.1 + 136 \cdot 15.8}{198.53 \cdot 21.539} \approx 0.9782$$

Problems

1. Consider

$$\bar{u}_1 = (1/2,\ 1/2,\ 1/2,\ 1/2),\ \bar{u}_2 = (1/2,\ 1/2,\ -1/2,\ -1/2),$$

$$\bar{u}_3 = (1/2,\ -1/2,\ 1/2,\ -1/2)$$

in $\mathbb{R}^4$.Can you find a vector $\bar{u}_4$ such that the vectors $\bar{u}_1$, $\bar{u}_2$, $\bar{u}_3$, $\bar{u}_4$ are orthonormal?

2. Let

$$W = span\,((1,2,3,4);\ (5,6,7,8))$$

Find a basis for $W^{\perp}$.

3. Let

$$V = Im \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Find $proj_V\, \bar{x}$, $\bar{x} = (1,3,1,7)$

Problems

4. Let $L = span\left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\right) \subset l^2$. Find the orthogonal projection of $(1, 0, 0, \ldots)$ onto L.

5. Among all the vectors in $\mathbb{R}^n$ whose components add up to 1, find the vector of minimal length.

6. Among all the unit vectors in $\mathbb{R}^n$, find the one for which the sum of the components is maximal.

7. There are three exams in your linear algebra class, and you theorize that your score in each exam (out of 100) will be numerically equal to the number of hours you study for that exam. The three exams count 20, 30, and 50, respectively, toward the final grade. If your (modest) goal is to score 76 in the course, how many hours a, b and c should you study for each of the three exams to minimize quantity $a^2 + b^2 + c^2$?

Orthonormal Basis

How to convert an arbitrary basis $\{\bar{v}_1, \ldots, \bar{v_m}\}$ of a linear subspace $V \subset \mathbb{R}^n$ into an orthonormal one, say $\{\bar{u}_1, \ldots, \bar{u_m}\}$?

Case1: dim $V = 1 \Rightarrow V$ is a line $\Rightarrow V = span\, \bar{v}_1 \Rightarrow \bar{u}_1 = \dfrac{\bar{v}_1}{|\bar{v}_1|}$

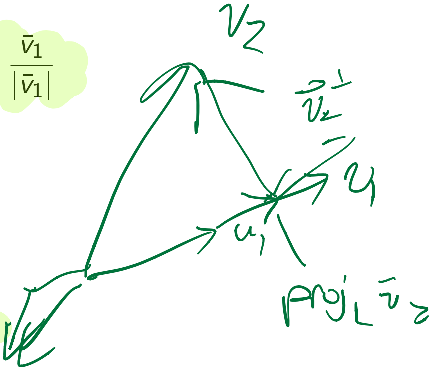
Case 2: dim $V = 2 \Rightarrow V$ is a plane $\Rightarrow V = span(\bar{v}_1,\, \bar{v}_2)$

$$\bar{u}_1 = \frac{\bar{v}_1}{|\bar{v}_1|} \quad \text{AND} \quad \bar{u}_2 \perp \bar{u}_1$$

$$\text{Let} \quad L = span\, \bar{u}_1 \Rightarrow \bar{u}_2 \perp proj_L \bar{v}_2$$

$$\bar{v}_2 = proj_L \bar{v}_2 + \underbrace{\bar{v}}_{\perp L} \Rightarrow \bar{u}_2 || \bar{v} = \bar{v}_2 - \underbrace{proj_L \bar{v}_2}_{=(\bar{u}_1, \bar{v}_2)\bar{u}_1}$$

Denote $\bar{v} = \bar{u}_2'$. Then $\bar{u}_2 = \dfrac{\bar{u}_2'}{|\bar{u}_2'|}$



Orthonormal Basis

Example:

$$V = span\left(\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\9\\9\\1\end{pmatrix}\right)$$

$$\bar{u}_1 = \begin{pmatrix}1/2\\1/2\\1/2\\1/2\end{pmatrix}, \bar{u}_2 = \begin{pmatrix}-1/2\\1/2\\1/2\\-1/2\end{pmatrix}$$

$$\bar{u}_2' = \begin{pmatrix}1\\9\\9\\1\end{pmatrix} - (1/2 \cdot 1 + 1/2 \cdot 9 + 1/2 \cdot 9 + 1/2 \cdot 1)\begin{pmatrix}1/2\\1/2\\1/2\\1/2\end{pmatrix}$$

Orthonormal Basis

Case 3: $\dim V = 3 \Rightarrow V = span(\bar{v}_1,\ \bar{v}_2,\ \bar{v}_3)$

$$\bar{u}_1 = \frac{\bar{v}_1}{|\bar{v}_1|}, \quad \bar{u}_2 = \frac{\bar{u}_2'}{|\bar{u}_2'|},\ \bar{u}_2' = \bar{v}_2 - (\bar{u}_1, \bar{v}_2)\bar{u}_1$$

Denote $E = span\,(\bar{u}_1,\ \bar{u}_2)$ and find $proj_E\, \bar{v}_3 = \bar{v}_3 - \underbrace{\bar{v}}_{\perp E}$

$$\bar{u}_3' = \bar{v} = \bar{v}_3 - proj_E\, \bar{v}_3 = \bar{v}_3 - (\bar{u}_1, \bar{v}_3)\bar{u}_1 - (\bar{u}_2, \bar{v}_3)\bar{u}_2$$

Gram-Schmidt Theorem

Theorem: Let $\{\bar{v}_1, \ldots, \bar{v_m}\}$ be a basis of a linear subspace $V \subset \mathbb{R}^n$. Since

$$\bar{v}_j = \bar{v}_j^{||} + \bar{v}_j^{\perp} \quad \forall j = 2, 3, \ldots$$

$$\bar{v}_j^{||} || span(\bar{v}_1, \ldots, \bar{v}_{j-1}) \text{ and } \bar{v}_j^{\perp} \perp span(\bar{v}_1, \ldots, \bar{v}_{j-1}),$$

so

$$\bar{u}_1 = \frac{\bar{v}_1}{|\bar{v}_1|}, \ \bar{u}_2 = \frac{\bar{v}_2^{\perp}}{|\bar{v}_2^{\perp}|}, \ldots, , \bar{u}_m = \frac{\bar{v}_m^{\perp}}{|\bar{v}_m^{\perp}|}$$

is the orthonormal basis of V.

$$\forall j > 2 \quad \bar{v}_j^{\perp} = \bar{v}_j - (\bar{u}_1, \bar{v}_j)\bar{u}_1 - \ldots - (\bar{u}_{j-1}, \bar{v}_j)\bar{u}_{j-1}$$

Gram-Schmidt Theorem: Remarks

Remark 1:

Using the Gram-Schmidt process, we change the basis $\mathfrak{B}_1 = \{\bar{v}_1, \ldots, \bar{v}_m\}$ to the basis $\mathfrak{B}_2 = \{\bar{u}_1, \ldots, \bar{u}_m\}$

$$\underbrace{(\bar{v}_1 \;\; \bar{v}_2 \; \ldots \; \bar{v}_m)}_{M} = \underbrace{(\bar{u}_1 \;\; \bar{u}_2 \; \ldots \; \bar{u}_m)}_{Q} \underbrace{R}_{\text{the change of basis matrix}}$$

$$QR\text{-factorization of } M: \quad M = QR$$

Remark 2: How to find the change of basis matrix R.

$\forall j > 2 \quad \bar{v}_j^{\perp} = \bar{v}_j - (\bar{u}_1, \bar{v}_j)\bar{u}_1 - \ldots - (\bar{u}_{j-1}, \bar{v}_j)\bar{u}_{j-1}$ and $\bar{v}_j^{\perp} = |\bar{v}_j^{\perp}|\bar{u}_j$

$$\bar{v}_j = \underbrace{(\bar{u}_1, \bar{v}_j)}_{r_{1j}} \bar{u}_1 + \underbrace{(\bar{u}_2, \bar{v}_j)}_{r_{2j}} \bar{u}_2 \ldots + \underbrace{(\bar{u}_{j-1}, \bar{v}_j)}_{r_{j-1,j}} \bar{u}_{j-1} + \underbrace{|\bar{v}_j^{\perp}|}_{r_{j,j}} \bar{u}_j$$

$$R = (r_{ij}) = \begin{cases} (\bar{u}_i, \bar{v}_j) & i < j \\ 0 & i > j, \\ |\bar{v}_j^{\perp}| & i = j,\ j > 2 \\ |\bar{v}_1| & i = j = 1 \end{cases}$$

QR-factorization. Example

Theorem (QR-factorization):
If the columns $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m$ of the matrix $M_{n\times m}$ are linearly independent, then there exists a matrix $Q_{n\times m}$ with orthogonal columns $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m$ and an upper triangular matrix R with positive diagonal entries such that

$$M = QR$$

Example 1:

$$A = \begin{pmatrix} 0 & -3 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}}_{R}$$

QR-factorization. Example

Example 2:

$$M = \begin{pmatrix} 2 & 2 \\ 1 & 7 \\ -2 & -8 \end{pmatrix}_{3\times 2} \rightarrow QR$$

$$\bar{v}_1 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \Rightarrow \bar{u}_1 = \begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \end{pmatrix}, \; \bar{u}_2' = \bar{v}_2 - (\bar{u}_1, \bar{v}_2)\bar{u}_1 =$$

$$= \begin{pmatrix} 2 \\ 7 \\ -8 \end{pmatrix} - \left(2\cdot\frac{2}{3} + 7\cdot\frac{1}{3} + 8\cdot\frac{2}{3}\right)\begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{pmatrix} =$$

$$= \begin{pmatrix} -4 \\ 4 \\ -2 \end{pmatrix} \Rightarrow \bar{u}_2 = \begin{pmatrix} -2/3 \\ 2/3 \\ -1/3 \end{pmatrix} \Rightarrow Q = \frac{1}{3}\begin{pmatrix} 2 & -2 \\ 1 & 2 \\ -2 & -1 \end{pmatrix}$$

u_2

$\tilde{u}_1$

QR-factorization. Example

Example 2 (cont):

$$M = \begin{pmatrix} 2 & 2 \\ 1 & 7 \\ -2 & -8 \end{pmatrix}_{3\times 2} \to QR, \quad Q = \frac{1}{3}\begin{pmatrix} 2 & -2 \\ 1 & 2 \\ -2 & -1 \end{pmatrix}$$

$$r_{11} = |\bar{v}_1| = 3, \quad r_{12} = (\bar{u}_1, \bar{v}_2) = 9, \quad r_{21} = 0, \quad r_{22} = |\bar{u}_2'| = 6$$

$$R = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$$

$$M = \begin{pmatrix} 2 & 2 \\ 1 & 7 \\ -2 & -8 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -2 \\ 1 & 2 \\ -2 & -1 \end{pmatrix}\begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$$