Vv417 Lecture 22

Jing Liu

UM-SJTU Joint Institute

November 21, 2019

- It is clearly better to use an orthonormal basis than using some other basis.
- So it is important to derive a process for constructing an orthonormal basis

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

for an n-dimensional inner product space $\mathcal{V}$ from an ordinary basis

$$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$$

- We want to construct the $\mathbf{u}_i$'s so that

$$\mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = \mathrm{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$$

- Our construct here is based on projections, to begin the process, let

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1 \implies \mathrm{span}(\mathbf{u}_1) = \mathrm{span}(\mathbf{x}_1)$$

- Let $\mathbf{p}_1$ denote the projection of $\mathbf{x}_2$ onto the subspace $\mathrm{span}(\mathbf{u}_1) = \mathrm{span}(\mathbf{x}_1)$

$$\mathbf{p}_1 = \langle \mathbf{x}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 \implies (\mathbf{x}_2 - \mathbf{p}_1) \perp \mathbf{u}_1$$

Q: Note that $\mathbf{x}_2 - \mathbf{p}_1 \neq \mathbf{0}$, why?

- Since $\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1$, thus

$$\mathbf{x}_2 - \mathbf{p}_1 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = \mathbf{x}_2 - \langle \mathbf{x}_2, \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1 \rangle \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1$$

- So if we set

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}_1\|}(\mathbf{x}_2 - \mathbf{p}_1)$$

  then $\mathbf{u}_2$ is a unit vector orthogonal to $\mathbf{u}_1$. It is clear that

$$\mathrm{span}(\mathbf{u}_1, \mathbf{u}_2) \subset \mathrm{span}(\mathbf{x}_1, \mathbf{x}_2)$$

- Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal, so they are linearly independent, and hence

$$\{\mathbf{u}_1, \mathbf{u}_2\}$$

  is an orthonormal basis for $\mathrm{span}(\mathbf{x}_1, \mathbf{x}_2)$, and

$$\mathrm{span}(\mathbf{u}_1, \mathbf{u}_2) = \mathrm{span}(\mathbf{x}_1, \mathbf{x}_2)$$

- To find $\mathbf{u}_3$, continue in the same way. Let $\mathbf{p}_2$ be the projection of $\mathbf{x}_3$ onto

$$\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2) = \operatorname{span}(\mathbf{x}_1, \mathbf{x}_2)$$

- Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is orthonormal,

$$\mathbf{p}_2 = \langle \mathbf{x}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{x}_3, \mathbf{u}_2 \rangle \mathbf{u}_2$$

- and if set

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3 - \mathbf{p}_2\|}(\mathbf{x}_3 - \mathbf{p}_2)$$

  then

$$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$$

  is an orthonormal basis for the subspace $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \operatorname{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$

- To obtain an orthonormal basis for the inner product space $\mathcal{V}$, we continue this process until we have n vectors in the set.

The Gram-Schmidt Process

Suppose $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a basis for the inner product space $\mathcal{V}$. Let

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1$$

and define $\mathbf{u}_2$, ..., $\mathbf{u}_n$ recursively by

$$\mathbf{u}_{k+1} = \left(\frac{1}{\|\mathbf{x}_{k+1} - \mathbf{p}_k\|}\right)(\mathbf{x}_{k+1} - \mathbf{p}_k) \qquad \text{for} \quad k = 1, \ldots, n-1$$

where

$$\mathbf{p}_k = \langle \mathbf{x}_{k+1}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{x}_{k+1}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{x}_{k+1}, \mathbf{u}_k \rangle \mathbf{u}_k$$

is the projection of $\mathbf{x}_{k+1}$ onto $\operatorname{span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$. Then the set

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

is an orthonormal basis for $\mathcal{V}$.

Exercise

Find an orthonormal basis for $\mathcal{P}_2$ if the inner product on $\mathcal{P}_2$ is defined by

$$\langle p, q\rangle = \sum_{i=1}^{3} p(x_i)q(x_i), \qquad \text{where } x_1 = -1,\ x_2 = 0, \text{ and } x_3 = 1.$$

Solution

- Starting with the basis $\{1, x, x^2\}$,

$$\|1\|^2 = \langle 1, 1\rangle = 1 + 1 + 1 = 3$$

- So

$$\mathbf{u}_1 = \left(\frac{1}{\|1\|}\right) 1 = \frac{1}{\sqrt{3}}$$

- Set

$$\mathbf{p}_1 = \langle x, \frac{1}{\sqrt{3}}\rangle = \left(-1 \cdot \frac{1}{\sqrt{3}} + 0 \cdot \frac{1}{\sqrt{3}} + 1 \cdot \frac{1}{\sqrt{3}}\right) = 0$$

Solution

- This means

$$x - \mathbf{p}_1 = x$$

$$\implies \|x - \mathbf{p}_1\|^2 = \langle x, x \rangle = (-1) \cdot (-1) + 0 \cdot 0 + 1 \cdot 1 = 2$$

- Hence

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} x$$

- Finally,

$$\mathbf{p}_2 = \langle x^2, \frac{1}{\sqrt{3}} \rangle \frac{1}{\sqrt{3}} + \langle x^2, \frac{x}{\sqrt{2}} \rangle \frac{x}{\sqrt{2}} = \frac{2}{3}$$

- thus

$$\mathbf{u}_3 = \frac{1}{\|x^2 - \mathbf{p}_2\|}(x^2 - \mathbf{p}_2) = \frac{\sqrt{6}}{2}(x^2 - \frac{2}{3})$$

Exercise

Find an orthonormal basis for $\operatorname{col}(\mathbf{A})$, where

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix}$$

Solution

- First of all, we have to determine the column space,

$$\mathbf{A} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

- Thus the columns are linearly independent, and

$$\operatorname{col}(\mathbf{A}) = \operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$$

Solution

- Apply Gram-Schmidt to the columns since they form a basis for $\mathrm{col}(\mathbf{A})$. Let

$$r_{11} = \|\mathbf{a}_1\| = \sqrt{\langle \mathbf{a}_1, \mathbf{a}_1 \rangle} = 2$$

- So the first vector in the orthonormal basis is

$$\mathbf{q}_1 = \frac{1}{r_{11}} \mathbf{a}_1$$

- To find the projection of $\mathbf{a}_2$ onto $\mathbf{q}_1$, we compute

$$r_{12} = \langle \mathbf{q}_1, \mathbf{a}_2 \rangle = 3$$

- To find the vector that is orthogonal to the span of $\mathbf{q}_1$,

$$\mathbf{a}_2 - r_{12}\mathbf{q}_1$$

Solution

- To normalize this orthogonal component of $\mathbf{a}_2$

$$r_{22} = \|\mathbf{a}_2 - r_{12}\mathbf{q}_1\| = 5$$

- Thus the second vector in the orthonormal basis is

$$\mathbf{q}_2 = \frac{1}{r_{22}}(\mathbf{a}_2 - r_{12}\mathbf{q}_1) = \frac{1}{r_{22}}(\mathbf{a}_2 - \frac{r_{12}}{r_{11}}\mathbf{a}_1)$$

- To find the projection of $\mathbf{a}_3$ onto $\mathcal{W} = \operatorname{span}\{\mathbf{q}_1, \mathbf{q}_2\}$,

$$\begin{aligned} r_{13} &= \langle \mathbf{q}_1, \mathbf{a}_3 \rangle = 2 \\ r_{23} &= \langle \mathbf{q}_2, \mathbf{a}_3 \rangle = -2 \end{aligned}$$

- To find the vector that is orthogonal to $\mathcal{W}$

$$\mathbf{a}_3 - r_{13}\mathbf{q}_1 - r_{23}\mathbf{q}_2$$

Solution

- To normalize this orthogonal component of $\mathbf{a}_3$,

$$r_{33} = \|\mathbf{a}_3 - r_{13}\mathbf{q}_1 - r_{23}\mathbf{q}_2| = 4$$

- Thus the last vector in the orthonormal basis is

$$\begin{aligned}\mathbf{q}_3 &= \frac{1}{r_{33}}\left(\mathbf{a}_3 - r_{13}\mathbf{q}_1 - r_{23}\mathbf{q}_2\right)\\ &= \frac{1}{r_{33}}\left(\mathbf{a}_3 - \frac{r_{13}}{r_{11}}\mathbf{a}_1 - \frac{r_{23}}{r_{22}}(\mathbf{a}_2 - \frac{r_{12}}{r_{11}}\mathbf{a}_1)\right)\end{aligned}$$

- Therefore the set is an orthonormal basis for $\operatorname{col}(\mathbf{A})$

$$\mathcal{B} = \{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} = \left\{ \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right\}$$

- If we retain all the inner products and norms computed during Gram-Schmidt process, a factorization of the matrix $\mathbf{A}$ can be obtained. For example,

$$\mathbf{Q} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3\end{bmatrix} = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33}\end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4\end{bmatrix}$$

$$\mathbf{QR} = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\end{bmatrix} \begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4\end{bmatrix} = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0\end{bmatrix} = \mathbf{A}$$

Gram-Schmidt QR Factorization

If $\mathbf{A}$ is an $m \times n$ matrix of rank n, then $\mathbf{A}$ can be factored into a product

$$\mathbf{A} = \mathbf{QR}$$

where $\mathbf{Q}$ is an $m \times n$ matrix with orthonormal column vectors and $\mathbf{R}$ is an upper triangular $n \times n$ matrix whose diagonal entries are all positive.

Proof

- Let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{n-1}$ be the projection vectors and $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ be the orthonormal basis of $\mathrm{col}(\mathbf{A})$ derived from the Gram-Schmidt process, and

$$r_{11} = \|\mathbf{a}_1\| \quad \text{and} \quad r_{kk} = \|\mathbf{a}_k - \mathbf{p}_{k-1}\| \qquad \text{for} \quad k = 2, \ldots, n$$
$$r_{ik} = \mathbf{q}_i^{\mathrm{T}}\mathbf{a}_k \qquad \text{for} \quad i = 1, \ldots, k-1 \quad \text{and} \quad k = i+1, \ldots, n$$

- By the Gram-Schmidt process,

$$\begin{aligned} r_{11}\mathbf{q}_1 &= \mathbf{a}_1 \\ r_{kk}\mathbf{q}_k &= \mathbf{a}_k - r_{1k}\mathbf{q}_1 - r_{2k}\mathbf{q}_2 - \cdots - r_{k-1,k}\mathbf{q}_{k-1} \qquad \text{for} \quad k = 2, \ldots, n \end{aligned}$$

- If we set $\mathbf{Q} = \begin{bmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{bmatrix}$ and defined $\mathbf{R} = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_{33} \end{bmatrix}$, then the kth column of the product $\mathbf{QR}$ will be $\mathbf{a}_k$ for $k = 1, \ldots, n$. Therefore,

$$\mathbf{QR} = \mathbf{A} \quad \square$$