Lecture 14 Outline

Wednesday, October 9, 2019 3:22 PM

Midterm Exam

Wed, October 16

In-class (~1 hr, 20 minutes)

Closed book, closed notes; no calculators, phones or computers

All material through Lecture 10 on linear transformations and HW 1-4

HW 4 Review

Orthogonal + Orthonormal Basis

Orthogonal & Orthonormal Basis

Sunday, October 6, 2019 10:42 PM

Recall that vectors are orthogonal, if

$$\underline{q}_0 \cdot \underline{q}_1 = 0$$

An orthogonal basis is one where all the vectors of the basis are orthogonal to each other

Example: $B = \{\underline{q}_0, \underline{q}_1, \underline{q}_2\}$, $\underline{q}_i \cdot \underline{q}_j = 0, i \neq j$

An orthonormal basis is one where

$$\underline{q}_i \cdot \underline{q}_j = 0 \quad i \neq j$$

$$\underline{q}_i \cdot \underline{q}_i = 1$$

Now, consider a matrix with orthonormal columns

$$\underline{Q} = \begin{bmatrix} | & | & | & & | \\ \underline{q}_1 & \underline{q}_2 & \underline{q}_3 & \cdots & \underline{q}_n \\ | & | & | & & | \end{bmatrix}$$

$$\underline{q}_1 \cdot \underline{q}_1 = 1 \qquad \underline{q}_1 \cdot \underline{q}_2 = 0$$

$$\vdots \qquad\qquad \vdots$$

$$\underline{q}_n \cdot \underline{q}_n = 1 \qquad \underline{q}_n \cdot \underline{q}_{n-1} = 0$$

Look at $\underline{Q}^T \underline{Q}$

Look at $\underline{Q}^T \underline{Q}$

$$\begin{bmatrix} \underline{q}_1^T \\ \underline{q}_2^T \\ \vdots \\ \underline{q}_n^T \end{bmatrix} \begin{bmatrix} \underline{q}_1 & \underline{q}_2 & \cdots & \underline{q}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & \ddots & 0 \\ & & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \underline{Q}^T \underline{Q} = \underline{I}$$

If $\underline{Q}$ is square, then $\underline{Q}\,\underline{Q}^T = \underline{I}$

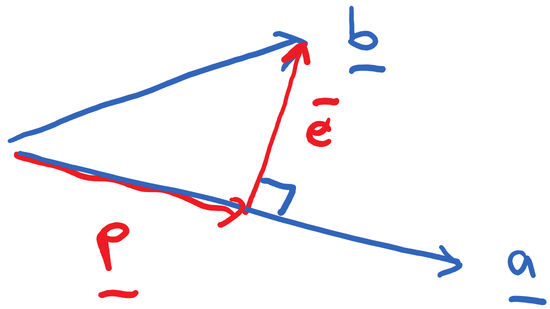
$\Rightarrow \underline{Q}^T = \underline{Q}^{-1}$ for square $\underline{Q}$ and is called a unitary matrix

How can $\underline{Q}$ be constructed?

Orthonormal Basis Construction

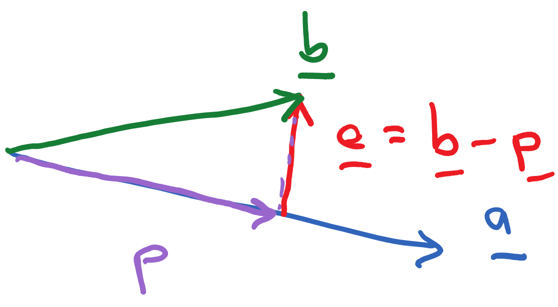
Given a set of vectors that span a subspace, find an orthonormal basis that also spans that subspace

Example: In 2D, let $\underline{a}$ & $\underline{b}$ be non-parallel vectors



$\underline{a}$ & $\underline{b}$ span $\mathbb{R}^2$, but are not orthogonal

Recall the projection of $\underline{b}$ onto $\underline{a}$



$$\underline{e} = \underline{b} - \underline{p} = \underline{b} - \frac{\underline{a}^T \underline{b}}{\underline{a}^T \underline{a}} \underline{a} \;;\; \underline{e} \perp \underline{a}$$

Thus, an orthonormal basis is

$$\underline{q}_1 = \frac{\underline{a}}{\|\underline{a}\|_2}, \quad \underline{q}_2 = \frac{\underline{e}}{\|\underline{e}\|_2} = \frac{\underline{b} - (\underline{a}^T \underline{b} / \underline{a}^T \underline{a}) \underline{a}}{\|e\|_2}$$

Are these vectors $\underline{q}_1$ + $\underline{q}_2$ unique?

No! For example, one could project $\underline{a}$ onto $\underline{b}$ instead

Now, consider a matrix and find an orthonormal basis to a column space $C(\underline{A})$, with

$$\underline{A} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{bmatrix} = \begin{bmatrix} \underline{a} & \underline{b} & \underline{c} \end{bmatrix}$$

Step 1: Set $\underline{t}_1 = \underline{a}$

Step 2: Project onto $\perp$ space of $\underline{a}$

$$\underline{t}_2 = \underline{b} - \frac{\underline{t}_1^T \underline{b}}{\underline{t}_1^T \underline{t}_1} \underline{t}_1 \qquad (\text{i.e. } \underline{I} - \underline{A})$$

$$\Rightarrow \underline{t}_1 \cdot \underline{t}_2 = 0$$

Step 3: Project onto $\perp$ space of $\underline{t}_1$ + $\underline{t}_2$

Step 3: Project onto $\perp$ space of $\underline{t}_1$ + $\underline{t}_2$

$$\underline{t}_3 = \underline{c} - \frac{\underline{t}_1^T \underline{c}}{\underline{t}_1^T \underline{t}_1} \underline{t}_1 - \frac{\underline{t}_2^T \underline{c}}{\underline{t}_2^T \underline{t}_2} \underline{t}_2$$

$$\Rightarrow \underline{t}_2 \cdot \underline{t}_3 = 0 \quad + \quad \underline{t}_1 \cdot \underline{t}_3 = 0$$

Step 4: Normalize

$$\underline{q}_1 = \frac{\underline{t}_1}{\|\underline{t}_1\|}, \quad \underline{q}_2 = \frac{\underline{t}_2}{\|\underline{t}_2\|}, \quad \underline{q}_3 = \frac{\underline{t}_3}{\|\underline{t}_3\|}$$

This proess is called Gram-Schmidt (G-S) orthonormalization

The result is an orthonormal basis to $C(\underline{A})$

How do $\underline{Q}$ and $\underline{A}$ relate?

Recall that G-S stated that

$$\underline{q}_1 = \frac{\underline{a}}{\|\underline{a}\|} = \frac{\underline{a}}{r_{11}} \Rightarrow \underline{a} = r_{11} \underline{q}_1$$

$$\underline{q}_2 = \frac{\underline{b} - r_{21}\underline{q}_1}{\|\underline{b} - r_{21}\underline{q}_1\|} = \frac{\underline{b} - r_{21}\underline{q}_1}{r_{22}} =$$

$$= \frac{1}{r_{22}}\underline{b} - \frac{r_{21}}{r_{22}}\underline{q}_1$$

$$\underline{b} = r_{21}\underline{q}_1 + r_{22}\underline{q}_2$$

Similarly, $\underline{c} = r_{31}\underline{q}_1 + r_{32}\underline{q}_2 + r_{33}\underline{q}_3$

$$\therefore \underline{A} = [\underline{a}\ \ \underline{b}\ \ \underline{c}] = \begin{bmatrix} \underline{q}_1 & \underline{q}_2 & \underline{q}_3 \end{bmatrix} \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ 0 & r_{22} & r_{32} \\ 0 & 0 & r_{33} \end{bmatrix}$$

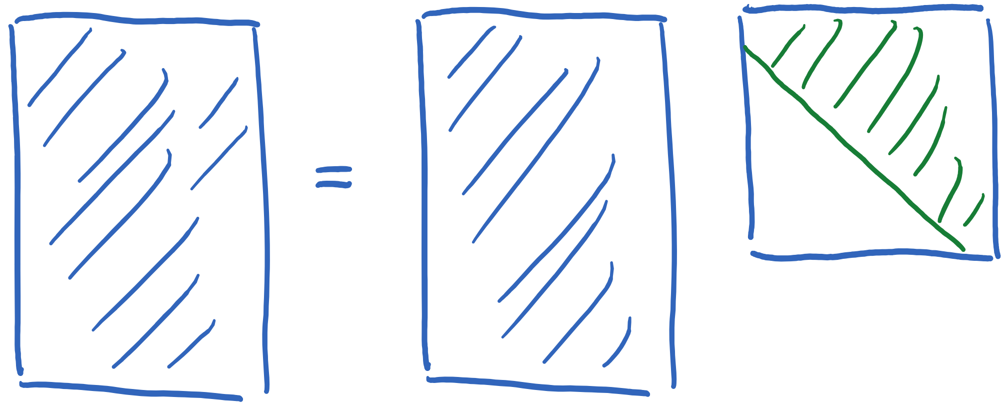
$$\underline{A} = \underbrace{\underline{Q}}_{\text{orthonormal matrix}}\ \underbrace{\underline{R}}_{\text{upper triangular}}$$

QR decomposition

This is actually called

Reduced QR Decomposition

typically written as $\underline{A} = \underline{\hat{Q}}\,\underline{\hat{R}}$

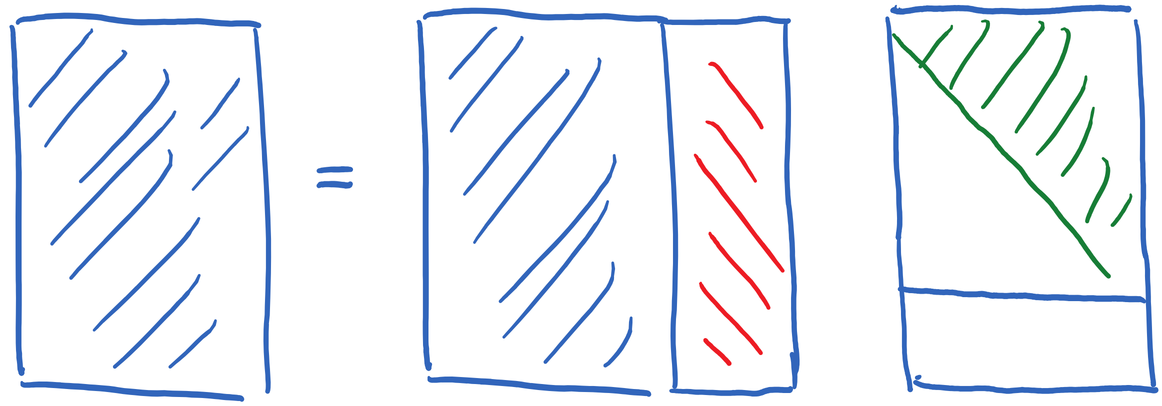


$$\underset{m \times n}{\underline{A}} \qquad \underset{m \times n}{\underline{\hat{Q}}} \qquad \underset{n \times n}{\underline{\hat{R}}}$$

Note: $$\underline{\hat{R}} = \begin{bmatrix} \underline{q}_1^T \underline{a} & \underline{q}_1^T \underline{b} & \underline{q}_1^T \underline{c} \\ 0 & \underline{q}_2^T b & \underline{q}_2^T \underline{c} \\ 0 & 0 & \underline{q}_3^T \underline{c} \end{bmatrix}$$

Also, one can determine a Full QR Factorization by appending colums to $\hat{Q}$ to make it $m \times m$

(typically with $m \geq n$)



$\underline{A}$ $\underline{Q}$ $\underline{R}$

$m \times n$ $m \times m$ $m \times n$

The columns $\underline{q}_j$ for $j > n$ must be orthogonal to the range$(\underline{A})$.

If $\text{rank}(\underline{A}) = n$, then these columns are the orthonormal basis to $\text{null}(\underline{A}^T)$

Why is this useful?

Theorem: Every $\underline{A} \in \mathbb{R}^{m \times n}$ with $m \geq n$ has a full QR factorization + a reduced

QR factorization

Theorem: Each $\underline{A} \in \mathbb{R}^{m \times n}$ with $m \geq n$ of full rank $(\text{rank}(\underline{A}) = n)$ has a unique reduced QR with $r_{jj} > 0$

All diagonals of $\underline{R}$ are positive

$\Rightarrow \underline{R}^{-1}$ exists and so does $\underline{R}^{-T}$

Now, let's look at $\underline{A}\,\underline{x} = \underline{b}$, $\underline{A}$ is full rank

Solve via $\underline{\hat{Q}}\,\underline{\hat{R}}$

① Decompose: $\underline{A} = \underline{\hat{Q}}\,\underline{\hat{R}}$ expensive

② $\underline{\hat{Q}}\,\underline{\hat{R}}\,\underline{x} = \underline{b} \Rightarrow \underline{\hat{Q}}^T \underline{\hat{Q}}\,\underline{\hat{R}}\,\underline{x} = \underline{\hat{Q}}^T \underline{b}$

③ $\underline{\hat{R}}\,\underline{x} = \underline{\hat{Q}}^T \underline{b}$ Solve (cheap)

If $\underline{b}$ changes (new right-hand side), but $\underline{A}$ does not, then it is cheap to solve for new solution $\underline{x}$

Note: Drop ^

Return to consider $\underline{A}^T\underline{A}\,\underline{x} = \underline{A}^T\underline{b}$

(Least squares formulation)

$\underline{A}^T\underline{A}\,\underline{x} = \underline{A}^T\underline{b}$ Find $\underline{A} = \underline{Q}\,\underline{R}$

$$(\underline{Q}\,\underline{R})^T(\underline{Q}\,\underline{R})\,\underline{x} = (\underline{Q}\,\underline{R})^T\underline{b}$$

$$\underline{R}^T\underline{Q}^T\underline{Q}\,\underline{R}\,\underline{x} = \underline{R}^T\underline{Q}^T\underline{b}$$

$$\underline{R}^T\underline{R}\,\underline{x} = \underline{R}^T\underline{Q}^T\underline{b}$$, where $\underline{R}^{-T}$ exists

$$\Rightarrow \underline{R}\,\underline{x} = \underline{Q}^T\underline{b}$$

If $m > n$ for $\underline{A} \in \mathbb{R}^{m \times n}$, then solving $\underline{R}\,\underline{x} = \underline{Q}^T \underline{b}$ is the solution that minimizes the error

Another advantage: Solving $\underline{R}\,\underline{x} = \underline{Q}^T \underline{b}$ is much more stable than

$$\underline{x} = (\underline{A}^T \underline{A})^{-1} \underline{A}^T \underline{b}$$

Remember condition number for this!

Classical Gram-Schmidt Algorithm

One algorithm for reduced QR of $\underline{A}$

Let $\underline{A} = [\underline{a}_1 \; \underline{a}_2 \; \ldots \; \underline{a}_n]$

Recall that $q_1 = \frac{\underline{a}}{} = \frac{\underline{a}}{}$

$- \|\underset{\sim}{q}\| \quad r_{11}$

$$\underset{\sim}{q}_2 = \frac{\underset{\sim}{a}_2 - r_{12}\underset{\sim}{q}_1}{r_{22}}$$

$\vdots$

where $r_{ij} = \underset{\sim}{q}_i^T \underset{\sim}{a}_j$ for $i \neq j$

and $|r_{jj}| = \| \underset{\sim}{a}_j - \sum_{i=1}^{j-1} r_{ij} \underset{\sim}{q}_i \|_2$

Note: r_{jj} can be either + or −, choose + value

Algorithm: Classical G-S

$$\underset{\sim}{Q} = \begin{bmatrix} \underset{\sim}{q}_1 & \underset{\sim}{q}_2 & \cdots & \underset{\sim}{q}_n \end{bmatrix}$$

for $j = 1:n$

$\underset{\sim}{v}_j = \underset{\sim}{a}_j$

for $i = 1:j-1$

$r_{ij} = q_i^T a_j$

$$r_{ij} = \underline{q}_i^T \underline{a}_j$$

$$\underline{v}_j = \underline{v}_j - r_{ij} \underline{q}_i$$

end

$$r_{jj} = \| \underline{v}_j \|_2$$

$$\underline{q}_j = \underline{v}_j / r_{jj}$$

end

Operation Count for G-S

Most expensive operation $r_{ij} = \underline{q}_i^T \underline{v}_j$

$+ \ \underline{v}_j = \underline{v}_j - r_{ij} \underline{q}_i$

$$\Rightarrow \sum_{i=1}^{n} \sum_{j=i+1}^{n} 4m \sim \sum_{i=1}^{n} 4mi \sim 2mn^2$$

However, Classical G-S is not numerically stable $\Rightarrow$ Roundoff errors cause issues

(We will not prove, because this requires complicated stability & error analysis. There will be a related HW problem.)

Consequently, a better method is needed:

Modified Gram-Schmidt