Matrix Decomposition

LU

QR

Spectral (square matrices)

SVD

Orthogonal & Orthonormal Basis

Recall that vectors are orthogonal, if

$$\underline{q}_0 \cdot \underline{q}_1 = 0 \qquad \text{or} \qquad \underline{q}_0^T \underline{q}_1 = 0$$

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

$$\underset{m \times n}{\underline{Q}} = \begin{bmatrix} | & | & | & & | \\ \underline{q}_1 & \underline{q}_2 & \underline{q}_3 & \cdots & \underline{q}_n \\ | & | & | & & | \end{bmatrix} \qquad \begin{matrix} \underline{q}_1 \cdot \underline{q}_1 = 1 & \underline{q}_1 \cdot \underline{q}_2 = 0 \\ \vdots & \vdots \\ \underline{q}_n \cdot \underline{q}_n = 1 & \underline{q}_n \cdot \underline{q}_{n-1} = 0 \end{matrix}$$

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

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

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

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

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

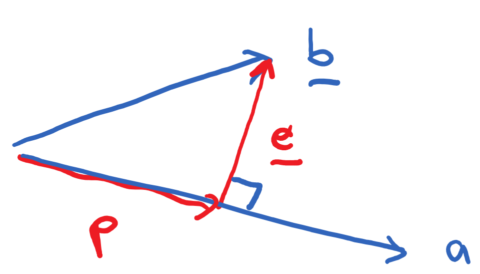
$$\Rightarrow \underline{Q}^T = \underline{Q}^{-1}$$ for square $\underline{Q}$ and is called a unitary matrix

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

Orthonormal Basis Construction

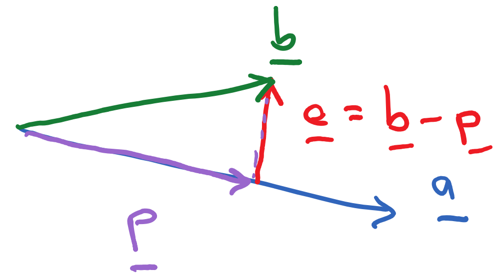
Given a set of vectors that span a subspace, find an orthonormal basis that also spans that subspace

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

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

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

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

Note: Columns are linearly independent

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

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

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

$$\Rightarrow \underset{\sim}{t}_1 \cdot \underset{\sim}{t}_2 = 0$$

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

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

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

Step 4: Normalize

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

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

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

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

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

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

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

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

$$\underline{A} = \underline{Q}\,\underline{R} \qquad QR$$

$\underline{A} = \underline{Q}\,\underline{R}$ QR decomposition

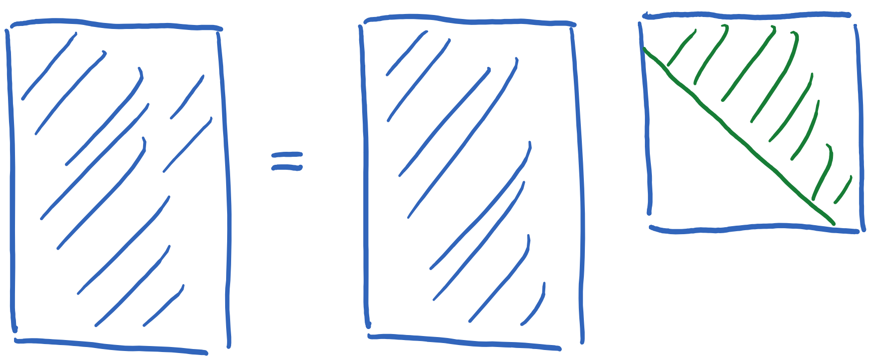
$\underline{Q}$: orthonormal matrix

$\underline{R}$: upper triangular

This is actually called

Reduced QR Decomposition

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

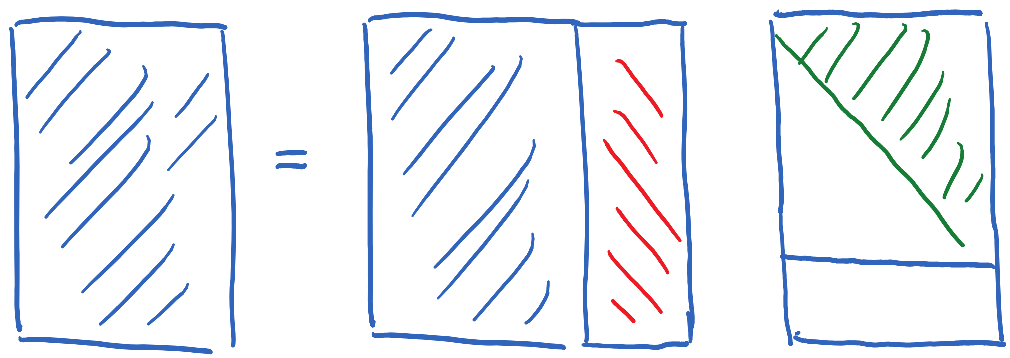


$\underline{A}$	$\underline{\hat{Q}}$	$\underline{\hat{R}}$
$m \times n$	$m \times n$	$n \times n$

$n \times n$: upper triangular (or right triangular matrix)

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

Also, one can determine a Full QR Factorization by appending colums to $\hat{Q}$ to make it $m \times m$ (typically with $m \geq n$)



$\underline{A}$ $m \times n$ = $\underline{Q}$ $m \times m$ $\underline{R}$ $m \times n$

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

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

nullspace

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 $\hat{\underline{R}}$ are positive

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

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

Solve via $\hat{\underline{Q}}\,\hat{\underline{R}}$ (Reduced QR decomposition)

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

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

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

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

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

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

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

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

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 $\underline{A} = [\underline{a}_1 \ \underline{a}_2 \ \dots \ \underline{a}_n]$

Recall that $\underline{q}_1 = \dfrac{\underline{a}_1}{\|\underline{a}_1\|} = \dfrac{\underline{a}_1}{r_{11}}$

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

$$\vdots$$

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

and $|r_{jj}| = \| \underline{a}_j - \sum^{j-1} r_{ij} \underline{q}_i \|$

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

Note: r_{jj} can be either + or −,

choose (+) value

Algorithm: Classical G-S

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

for $j = 1:n$ (loop over columns j)

$\quad \underline{v}_j = \underline{a}_j$

$\quad$ for $i = 1:j-1$ (loop over previous columns)

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

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

$\quad$ end

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

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

end

Operation Count for G-S

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

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

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

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

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

Consequently, a better method is needed:

Modified Gram-Schmidt

How to compute QR factorization?

$\underset{m\times n}{\underline{A}} = \underset{m\times n}{\underline{\hat{Q}}}\, \underset{n\times n}{\underline{\hat{R}}}$ $\leftrightarrow$ Partial (Reduced) QR factorization ($\underline{\hat{R}}$ is square)

$\underset{m\times n}{\underline{A}} = \underset{m\times m}{\underline{Q}}\, \underset{m\times n}{\underline{R}}$ $\leftrightarrow$ Full QR factorization ($\underline{Q}$ is square)

$$\underline{\hat{Q}}^T \underline{\hat{Q}} = \underset{n\times n}{\underline{I}}$$

$$\underline{Q}^T \underline{Q} = \underline{Q}\,\underline{Q}^T = \underline{I} \Rightarrow \underline{Q}^T = \underline{Q}^{-1}$$ (unitary $\underline{Q}$)

Classical Gram-Schmidt Algorithm →

projection based, not stable numerically (round off error)

Modified Gram-Schmidt

Recall that projection can be written as a matrix-vector product ⇒

$$\underline{q}_1 = \frac{\underline{P}_1 \underline{a}_1}{\|\underline{P}_1 \underline{a}_1\|}, \quad \underline{q}_2 = \frac{\underline{P}_2 \underline{a}_2}{\|\underline{P}_2 \underline{a}_2\|}, \text{ etc.}$$

for some $\underline{P}_j$

Let $\hat{\underline{Q}}_{j-1}$ be the $m \times (j-1)$ matrix of the first $j-1$ columns of $\hat{\underline{Q}}$

where

$$\hat{\underline{Q}} = [\underline{q}_1 \ \underline{q}_2 \ \cdots \ \underline{q}_n]$$

$$\hat{\underline{Q}}_{j-1} = [\underline{q}_1 \ \underline{q}_2 \ \cdots \ \underline{q}_{j-1}]$$

Then

$$\underline{P}_j = \underline{I} - \hat{\underline{Q}}_{j-1} \hat{\underline{Q}}_{j-1}^T \rightarrow \text{matrices of}$$

the form $\underline{I} - \underline{v}\,\underline{v}^T$ project onto the perpendicular space of $\underline{v}$

Thus, $\underline{P}_j$ is nothing but the repeated perpendicular projections of each prior vector in $\hat{\underline{Q}}$ or

$$\underline{P}_j = \underline{P}_{\perp \underline{q}_{j-1}} \underline{P}_{\perp \underline{q}_{j-2}} \cdots \underline{P}_{\perp \underline{q}_2} \underline{P}_{\perp \underline{q}_1}$$

with $\underset{\sim}{P}_1 = \underline{I}$

Each $\underset{\sim}{P}_{\perp q_j}$ projects onto the space perpendicular to $\underset{\sim}{q}_j$

Modified Gram-Schmidt uses these ideas to reverse the order of operations, such that

Algorithm: Modified G-S

for $i = 1:n$

$\quad \underset{\sim}{v}_i = \underset{\sim}{q}_i$

end

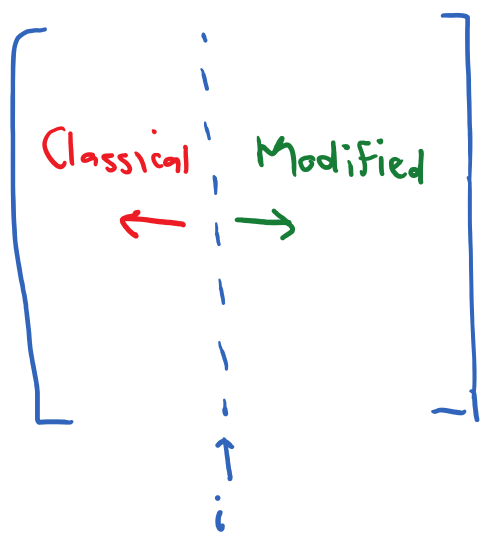
for $i = 1:n$ (loop over columns i)

$\quad r_{ii} = \| \underset{\sim}{v}_i \|$

$\quad \underset{\sim}{q}_i = \underset{\sim}{v}_i / r_{ii}$

$\quad$ for $j = i+1:n$ (loop over following columns)

$\quad\quad r_{ij} = \underset{\sim}{q}_i^T \underset{\sim}{v}_j$



Reduces effects of roundoff error

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

end

end

Operation count for Modified G-S is identical to Classical G-S: $O(2mn^2)$

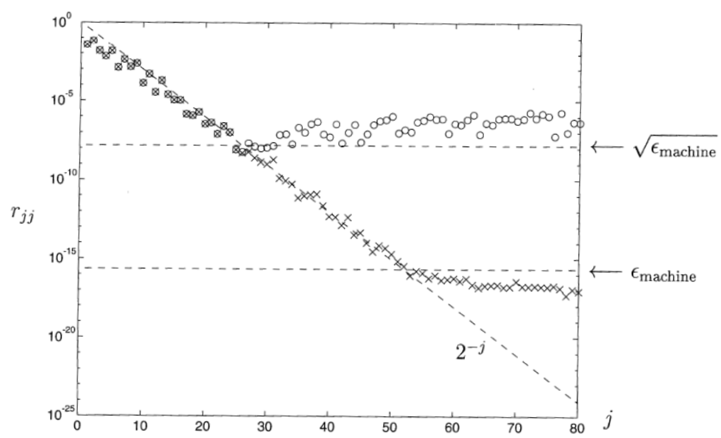


Figure 9.1. *Computed r_{jj} versus j for the QR factorization of a matrix with exponentially graded singular values. On this computer with about 16 digits of relative accuracy, the classical Gram–Schmidt algorithm produces the numbers represented by circles and the modified Gram–Schmidt algorithm produces the numbers represented by crosses.*

Trefethen & Bau (1997)

Householder Triangularization

Look at G-S again

In G-S, each operation to compute a

column of $\hat{\underline{Q}}$ is an upper triangular matrix multiplication

$$\underline{A}\,\underbrace{\underline{R}_1\underline{R}_2\cdots\underline{R}_n}_{\hat{\underline{R}}^{-1}} = \hat{\underline{Q}} \Rightarrow \underline{A} = \hat{Q}\hat{R}$$

This is called Triangular Orthogonalization:

R gives Q

One can do the reverse: repeated applications of Q give R

$$\underbrace{\underline{Q}_n\underline{Q}_{n-1}\cdots\underline{Q}_2\underline{Q}_1}_{\hat{\underline{Q}}^{T}}\,\underline{A} = \hat{\underline{R}} \Rightarrow \underline{A} = \hat{\underline{Q}}\hat{\underline{R}}$$

This is called Orthogonal Triangularization

Q gives R