Introduction to Simulation - Lecture 5

# QR Factorization

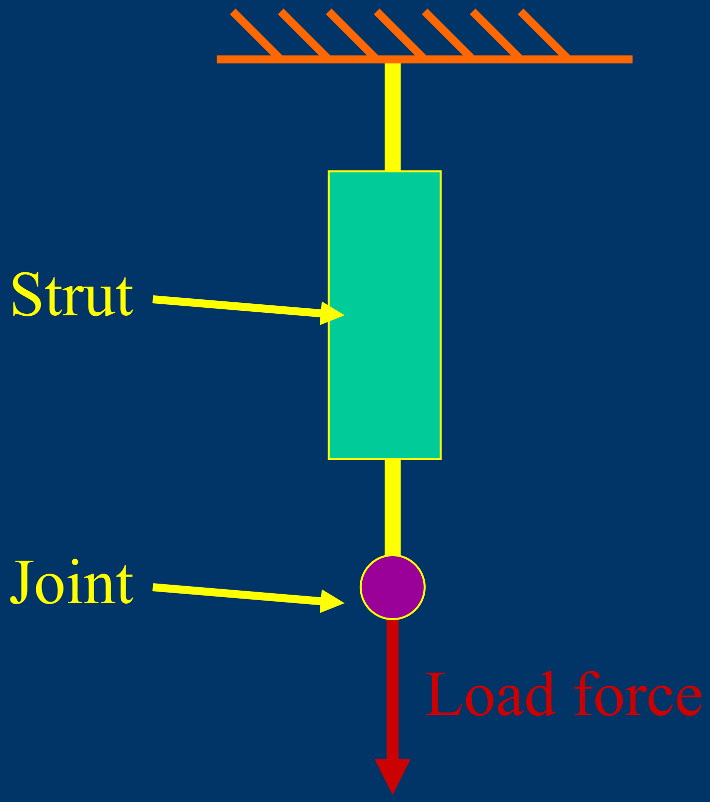
Jacob White

Thanks to Deepak Ramaswamy, Michal Rewienski, and Karen Veroy

# QR Factorization

## Singular Example

### LU Factorization Fails

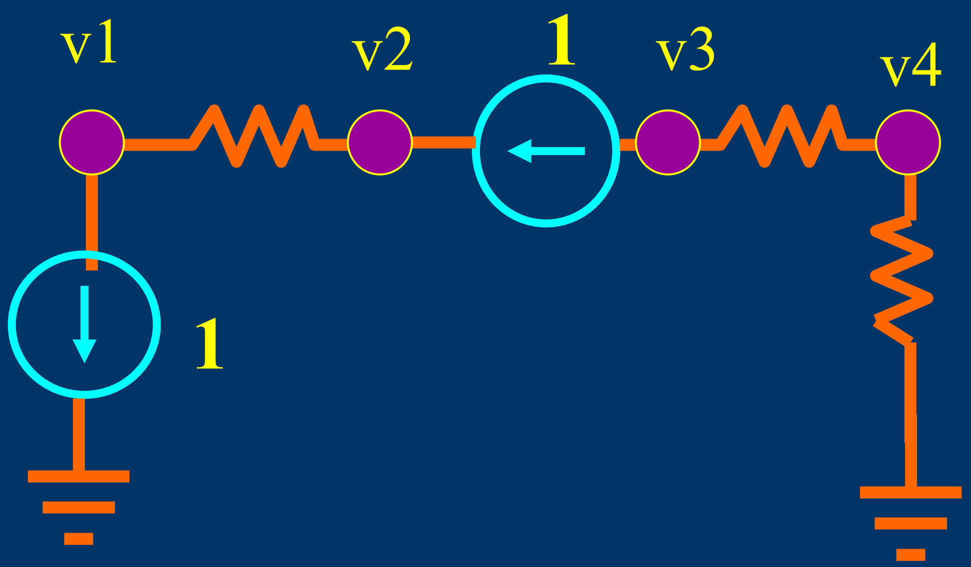



The resulting nodal matrix is SINGULAR, but a solution exists!

# QR Factorization

## Singular Example

### LU Factorization Fails




The resulting nodal matrix is SINGULAR, but a solution exists!

# QR Factorization

## Singular Example

Recall weighted sum of columns view of systems of equations

$$\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \vec{M}_1 & \vec{M}_2 & \cdots & \vec{M}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

$$x_1\vec{M}_1 + x_2\vec{M}_2 + \cdots + x_N\vec{M}_N = b$$

**M is singular but b is in the span of the columns of M**

# QR Factorization

## Orthogonalization

### If M has orthogonal columns

Orthogonal columns implies:

$$\vec{M}_i \bullet \vec{M}_j = 0 \qquad i \neq j$$

Multiplying the weighted columns equation by ith column:

$$\vec{M}_i \bullet \left( x_1\vec{M}_1 + x_2\vec{M}_2 + \cdots + x_N\vec{M}_N \right) = \vec{M}_i \bullet b$$

Simplifying using orthogonality:

$$x_i \left( \vec{M}_i \bullet \vec{M}_i \right) = \vec{M}_i \bullet b \quad \Rightarrow \quad x_i = \frac{\vec{M}_i \bullet b}{\left( \vec{M}_i \bullet \vec{M}_i \right)}$$

# QR Factorization

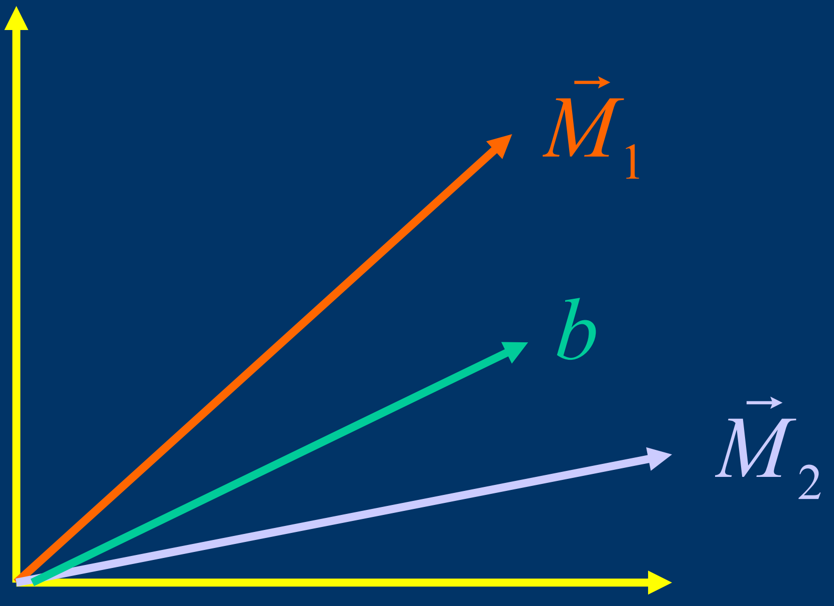
## Orthogonalization

### Orthonormal M - Picture

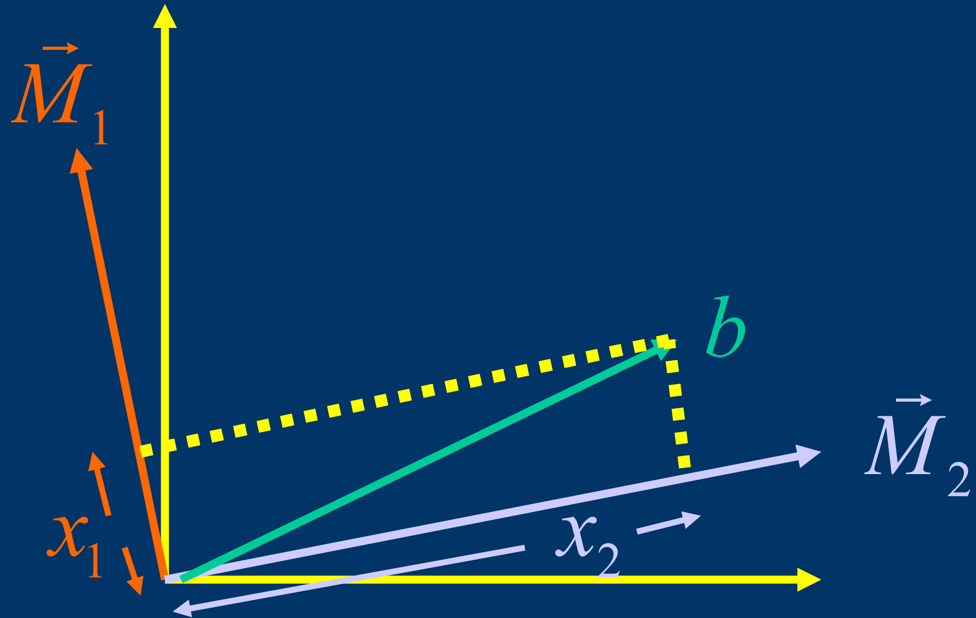
M is orthonormal if:

$$\vec{M}_i \bullet \vec{M}_j = 0 \quad i \neq j \quad \text{and} \quad \vec{M}_i \bullet \vec{M}_i = 1$$

Picture for the two-dimensional case




Non-orthogonal Case




Orthogonal Case

# QR Factorization

## Orthogonalization

### QR Algorithm Key Idea

$$\underbrace{\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \vec{M}_1 & \vec{M}_2 & \cdots & \vec{M}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}}_{Original\ Matrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \Rightarrow \underbrace{\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \vec{Q}_1 & \vec{Q}_2 & \cdots & \vec{Q}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}}_{Matrix\ with\ Orthonormal\ Columns} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

$$Qy = b \quad \Rightarrow \quad y = Q^T b$$

How to perform the conversion?

# QR Factorization

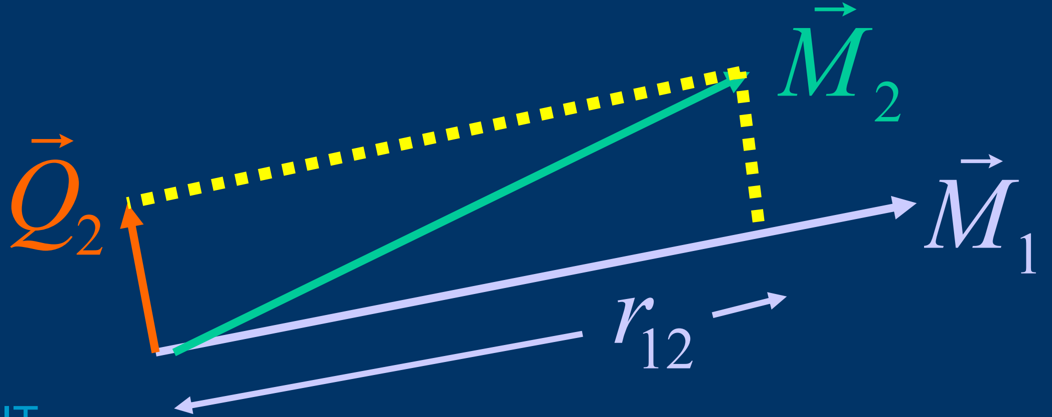
## Orthogonalization

### Projection Formula

Given $\vec{M}_1, \ \vec{M}_2,$ find $\vec{Q}_2 = \vec{M}_2 - r_{12}\vec{M}_1$ so that

$$\vec{M}_1 \bullet \vec{Q}_2 = \vec{M}_1 \bullet \left( \vec{M}_2 - r_{12}\vec{M}_1 \right) = 0$$

$$\Rightarrow \quad r_{12} = \frac{\vec{M}_1 \bullet \vec{M}_2}{\vec{M}_1 \bullet \vec{M}_1}$$



# QR Factorization

Orthogonalization

Normalization

Formulas simplify if we normalize

$$\vec{Q}_1 = \frac{1}{\sqrt{\vec{M}_1 \bullet \vec{M}_1}} \vec{M}_1 = \frac{1}{r_{11}} \vec{M}_1 \Rightarrow \vec{Q}_1 \bullet \vec{Q}_1 = 1$$

Now find $\tilde{\vec{Q}}_2 = \vec{M}_2 - r_{12}\vec{Q}_1$ so that $\tilde{\vec{Q}}_2 \bullet \vec{Q}_1 = 0$

$$\Rightarrow \quad r_{12} = \vec{Q}_1 \bullet \vec{M}_2$$

Finally $\vec{Q}_2 = \frac{1}{\sqrt{\tilde{\vec{Q}}_2 \bullet \tilde{\vec{Q}}_2}} \tilde{\vec{Q}}_2 = \frac{1}{r_{22}} \tilde{\vec{Q}}_2$

# QR Factorization

## Orthogonalization

### How was a 2x2 matrix converted?

Since Mx should equal Qy, we can relate x to y

$$\begin{bmatrix} \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \vec{M}_1 + x_2 \vec{M}_2 = \begin{bmatrix} \uparrow & \uparrow \\ \vec{Q}_1 & \vec{Q}_2 \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y_1 \vec{Q}_1 + y_2 \vec{Q}_2$$

$$\vec{M}_1 = r_{11} \vec{Q}_1 \qquad \vec{M}_2 = r_{22} \vec{Q}_2 + r_{12} \vec{Q}_1$$

$$\Rightarrow \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{bmatrix} \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \uparrow & \uparrow \\ \vec{Q}_1 & \vec{Q}_2 \\ \downarrow & \downarrow \end{bmatrix}}_{Orthonormal} \underbrace{\begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}}_{Upper\ Triangular} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Two Step Solve Given QR

Step 1) $QRx = b \Rightarrow Rx = Q^T b = \tilde{b}$

Step 2) Backsolve $Rx = \tilde{b}$

# QR Factorization

## Orthogonalization

### The General Case

# 3x3 Case

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 & \vec{M}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \Rightarrow \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 - r_{12}\vec{M}_1 & \vec{M}_3 - r_{13}\vec{M}_1 - r_{23}\vec{M}_2 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$$

To Insure the third column is orthogonal

$$\vec{M}_1 \bullet \left( \vec{M}_3 - r_{13}\vec{M}_1 - r_{23}\vec{M}_2 \right) = 0$$

$$\vec{M}_2 \bullet \left( \vec{M}_3 - r_{13}\vec{M}_1 - r_{23}\vec{M}_2 \right) = 0$$

# QR Factorization

## Orthogonalization

### Must Solve Equations for Coefficients in 3x3 Case

$$\vec{M}_1 \bullet \left( \vec{M}_3 - r_{13}\vec{M}_1 - r_{23}\vec{M}_2 \right) = 0$$

$$\vec{M}_2 \bullet \left( \vec{M}_3 - r_{13}\vec{M}_1 - r_{23}\vec{M}_2 \right) = 0$$

$$\Downarrow$$

$$\begin{bmatrix} \vec{M}_1 \bullet \vec{M}_1 & \vec{M}_1 \bullet \vec{M}_2 \\ \vec{M}_2 \bullet \vec{M}_1 & \vec{M}_2 \bullet \vec{M}_2 \end{bmatrix} \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} = \begin{bmatrix} \vec{M}_1 \bullet \vec{M}_3 \\ \vec{M}_2 \bullet \vec{M}_3 \end{bmatrix}$$

# QR Factorization

## Orthogonalization

### Must Solve Equations for Coefficients

To Orthogonalize the Nth Vector

$$\begin{bmatrix} \vec{M}_1 \bullet \vec{M}_1 & \cdots & \vec{M}_1 \bullet \vec{M}_{N-1} \\ \vdots & \ddots & \vdots \\ \vec{M}_{N-1} \bullet \vec{M}_1 & \cdots & \vec{M}_{N-1} \bullet \vec{M}_{N-1} \end{bmatrix} \begin{bmatrix} r_{1,N} \\ \vdots \\ r_{N-1,N} \end{bmatrix} = \begin{bmatrix} \vec{M}_1 \bullet \vec{M}_N \\ \vdots \\ \vec{M}_{N-1} \bullet \vec{M}_N \end{bmatrix}$$

$N^2$ inner products requires $N^3$ work

# QR Factorization

## Orthogonalization

### Use previously orthogonalized vectors

## 3x3 Case

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 & \vec{M}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \Rightarrow \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{M}_1 & \vec{M}_2 - r_{12}\vec{Q}_1 & \vec{M}_3 - r_{13}\vec{Q}_1 - r_{23}\vec{Q}_2 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$$

To Insure the third column is orthogonal

$$\vec{Q}_1 \bullet \left( \vec{M}_3 - \vec{Q}_1 r_{13} - \vec{Q}_2 r_{23} \right) = 0 \Rightarrow r_{13} = \vec{Q}_1 \bullet \vec{M}_3$$

$$\vec{Q}_2 \bullet \left( \vec{M}_3 - \vec{Q}_1 r_{13} - \vec{Q}_2 r_{23} \right) = 0 \Rightarrow r_{23} = \vec{Q}_2 \bullet \vec{M}_3$$

# QR Factorization

## Basic Algorithm

### "Modified Gram-Schmidt"

For i = 1 to N "For each Source Column"

$$r_{ii} = \sqrt{\vec{M}_i \bullet \vec{M}_i}$$

$$\vec{Q}_i = \frac{1}{r_{ii}} \vec{M}_i$$

Normalize

$$\sum_{i=1}^{N} 2N \approx 2N^2 \text{ operations}$$

For j = i+1 to N { "For each target Column right of source"

$$r_{ij} \leftarrow \vec{M}_j \bullet \vec{Q}_i$$

$$\vec{M}_j \Leftarrow \vec{M}_j - r_{ij}\vec{Q}_i$$

$$\sum_{i=1}^{N} (N-i)2N \approx N^3 \text{ operations}$$

# QR Factorization

## Basic Algorithm

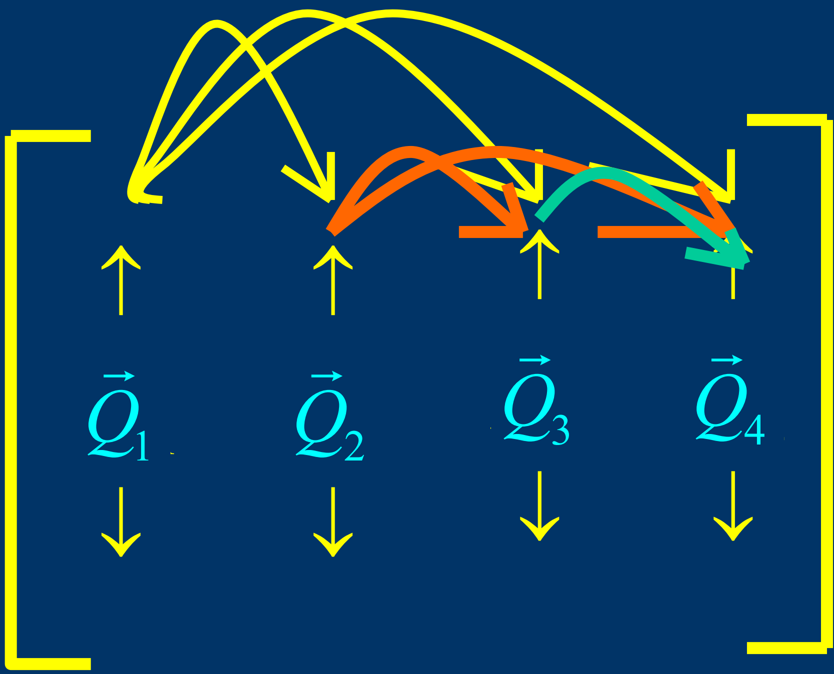
### "By Picture"

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow & \cdots & \uparrow \\ \vec{Q}_1 & \vec{Q}_2 & \vec{Q}_3 & \cdots & \vec{Q}_N \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1N} \\ 0 & r_{22} & r_{23} & \cdots & r_{2N} \\ 0 & 0 & r_{33} & \cdots & r_{3N} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{NN} \end{bmatrix}$$

# QR Factorization

## Basic Algorithm

### "By Picture"



$$
\begin{bmatrix} \vec{Q}_1 & \vec{Q}_2 & \vec{Q}_3 & \vec{Q}_4 \end{bmatrix}
\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ & r_{22} & r_{23} & r_{24} \\ & & r_{33} & r_{34} \\ & & & r_{44} \end{bmatrix}
$$

# QR Factorization

## Basic Algorithm

### Zero Column

What if a Column becomes Zero?

$$\begin{bmatrix} \uparrow & 0 & \uparrow & \cdots & \uparrow \\ \vec{Q}_1 & 0 & \tilde{\vec{M}}_3 & \cdots & \tilde{\vec{M}}_N \\ \downarrow & 0 & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1N} \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Matrix MUST BE Singular!

1) Do not try to normalize the column.
2) Do not use the column as a source for orthogonalization.
3) Perform backward substitution as well as possible

# QR Factorization

## Basic Algorithm

### Zero Column Continued

Resulting QR Factorization

$$\begin{bmatrix} \uparrow & 0 & \uparrow & \cdots & \uparrow \\ \vec{Q}_1 & 0 & \vec{Q}_3 & \cdots & \vec{Q}_N \\ \downarrow & 0 & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1N} \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & r_{33} & \cdots & r_{3N} \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{NN} \end{bmatrix}$$

# QR Factorization

## Singular Example

Recall weighted sum of columns view of systems of equations

$$\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \vec{M}_1 & \vec{M}_2 & \cdots & \vec{M}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \qquad x_1\vec{M}_1 + x_2\vec{M}_2 + \cdots + x_N\vec{M}_N = b$$

Two Cases when M is singular

Case 1) $b \in span\{\vec{M}_1,..,\vec{M}_N\} \Rightarrow b \in span\{\vec{Q}_1,..,\vec{Q}_N\}$

Case 2) $b \notin span\{\vec{M}_1,..,\vec{M}_N\}$, How accurate is $x$?

# QR Factorization

## Minimization View

### Alternative Formulations

Definition of the Residual R: $R(x) \equiv b - Mx$

| Find x which satisfies | Minimize over all x |
|---|---|
| $Mx = b$ | $R(x)^T R(x) = \sum_{i=1}^{N} \left(R_i(x)\right)^2$ |

Equivalent if $b \in span\{cols(M)\}$

$$\Rightarrow Mx = b \text{ and } \min_x R(x)^T R(x) = 0$$

Minimization extends to non-singular or nonsquare case!

# QR Factorization

## Minimization View

### One-dimensional Minimization

Suppose $x = x_1\vec{e}_1$ and therefore $Mx = x_1 M\vec{e}_1 = x_1\vec{M}_1$

<u>One dimensional Minimization</u>

$$R(x)^T R(x) = \left(b - x_1 M\vec{e}_1\right)^T \left(b - x_1 M\vec{e}_1\right)$$

$$= b^T b - 2x_1 b^T M\vec{e}_1 + x_1^2 \left(M\vec{e}_1\right)^T \left(M\vec{e}_1\right)$$

$$\frac{d}{dx} R(x)^T R(x) = -2b^T M\vec{e}_1 + 2x_1 \left(M\vec{e}_1\right)^T \left(M\vec{e}_1\right) = 0$$

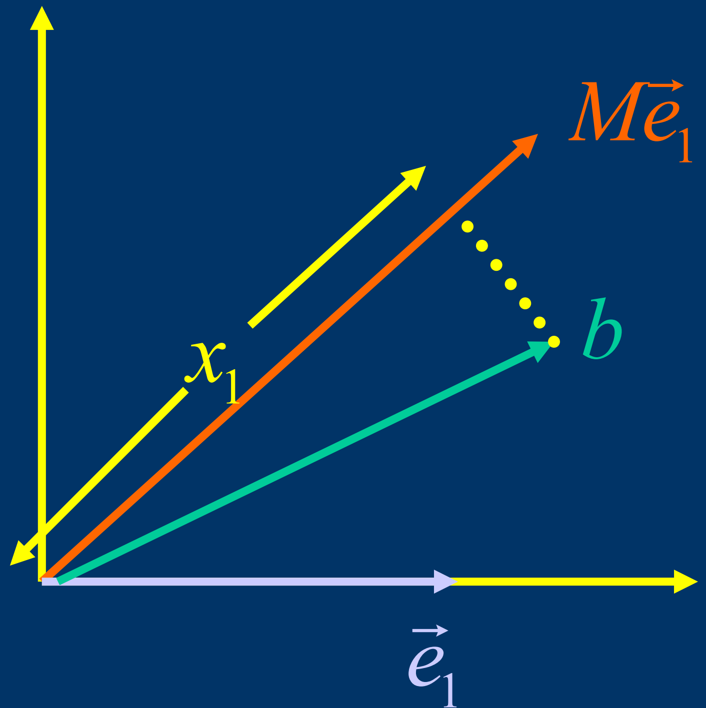
$$x_1 = \frac{b^T M\vec{e}_1}{\vec{e}_1^T M^T M\vec{e}_1}$$

Normalization

# QR Factorization

## Minimization View

### One-dimensional Minimization, Picture




$$x_1 = \frac{b^T M\vec{e}_1}{\vec{e}_1^T M^T M\vec{e}_1}$$

One dimensional minimization yields same result as projection on the column!

# QR Factorization

## Minimization View

### Two-dimensional Minimization

Now $x = x_1\vec{e}_1 + x_2\vec{e}_2$ and $Mx = x_1M\vec{e}_1 + x_2M\vec{e}_2$

Residual Minimization

$$R(x)^T R(x) = \left(b - x_1M\vec{e}_1 - x_2M\vec{e}_2\right)^T \left(b - x_1M\vec{e}_1 - x_2M\vec{e}_2\right)$$

$$= b^T b - 2x_1 b^T M\vec{e}_1 + x_1^2 \left(M\vec{e}_1\right)^T \left(M\vec{e}_1\right)$$

$$-2x_2 b^T M\vec{e}_2 + x_2^2 \left(M\vec{e}_2\right)^T \left(M\vec{e}_2\right)$$

Coupling Term

$$+2x_1x_2 \left(M\vec{e}_1\right)^T \left(M\vec{e}_2\right)$$

## More General Search Directions

$$x = v_1 \vec{p}_1 + v_2 \vec{p}_2 \text{ and } Mx = v_1 M\vec{p}_1 + v_2 M\vec{p}_2$$

$$\text{span}\{\vec{p}_1, \vec{p}_2\} = \text{span}\{\vec{e}_1, \vec{e}_2\}$$

$$R(x)^T R(x) = b^T b - 2v_1 b^T M\vec{p}_1 + v_1^2 (M\vec{p}_1)^T (M\vec{p}_1)$$

$$-2v_2 b^T M\vec{p}_2 + v_2^2 (M\vec{p}_2)^T (M\vec{p}_2)$$

$$+2v_1 v_2 (M\vec{p}_1)^T (M\vec{p}_2)$$

Coupling Term

If $\vec{p}_1^{\,T} M^T M\vec{p}_2 = 0$ Minimizations Decouple!!

# QR Factorization

## Minimization View

### Forming $M^TM$ orthogonal Minimization Directions

ith search direction equals $\mathbf{M^TM}$ orthogonalized unit vector

$$\vec{p}_i = \vec{e}_i - \sum_{j=1}^{i-1} r_{ji}\vec{p}_j \qquad \vec{p}_i^{\,T} M^T M \vec{p}_j = 0$$

Use previous orthogonalized Search directions

$$\Rightarrow r_{ji} = \frac{\left(M\vec{p}_j\right)^T \left(M\vec{e}_i\right)}{\left(M\vec{p}_j\right)^T \left(M\vec{p}_j\right)}$$

# QR Factorization

## Minimization View

### Minimizing in the Search Direction

Decoupled minimizations done individually

Minimize: $v_i^2 \left(M\vec{p}_i\right)^T \left(M\vec{p}_i\right) - 2v_i b^T M\vec{p}_i$

Differentiating: $2v_i \left(M\vec{p}_i\right)^T \left(M\vec{p}_i\right) - 2b^T M\vec{p}_i = 0$

$$\Rightarrow \quad v_i = \frac{b^T M\vec{p}_i}{\left(M\vec{p}_i\right)^T \left(M\vec{p}_i\right)}$$

# QR Factorization

Minimization View

**Minimization Algorithm**

For i = 1 to N "For each Target Column"

$$\vec{p}_i = \vec{e}_i$$

For j = 1 to i-1 "For each Source Column left of target"

$$\left.\begin{aligned} r_{ij} &\leftarrow \vec{p}_j^T M^T M \vec{p}_i \\ \vec{p}_i &\Leftarrow \vec{p}_i - r_{ij} \vec{p}_j \end{aligned}\right\}$$

Orthogonalize Search Direction

$$\left.\begin{aligned} r_{ii} &= \sqrt{M\vec{p}_i \bullet M\vec{p}_i} \\ \vec{p}_i &\Leftarrow \frac{1}{r_{ii}} \vec{p}_i \end{aligned}\right\}$$

Normalize search direction

$$x = x + v_i \vec{p}_i$$

# QR Factorization

## Minimization and QR

### Comparison

$\vec{Q}_1$ $\quad \vec{Q}_2 \quad \cdots\cdots \quad \vec{Q}_N$ } Orthonormal

M $\qquad$ M $\qquad$ M

$$\underbrace{\frac{1}{r_{11}}\vec{e}_1}_{\vec{p}_1} \quad \underbrace{\frac{1}{r_{22}}\left(\vec{e}_2 - r_{12}\vec{e}_1\right)}_{\vec{p}_2} \quad \cdots\cdots \quad \underbrace{\frac{1}{r_{NN}}\left(\vec{e}_2 - \sum r_{iN}\vec{e}_i\right)}_{\vec{p}_N}$$

} $M^TM$ Orthonormal

# QR Factorization

## Search Direction

Orthogonalized unit vectors → search directions

$$\underbrace{\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_N\}}_{\text{Unit Vectors}} \xrightarrow[\text{Orthogonalization}]{M^TM} \underbrace{\{\vec{p}_1, \ldots, \vec{p}_N\}}_{\text{Search Directions}}$$

Could use other sets of starting vectors

$$\underbrace{\{b, Mb, M^2b, \ldots\}}_{\text{Krylov-Subspace}} \xrightarrow[\text{Orthogonalization}]{M^TM} \underbrace{\{\vec{p}_1, \ldots, \vec{p}_N\}}_{\text{Search Directions}}$$

**Why?**

# Summary

- QR Algorithm
  - Projection Formulas
  - Orthonormalizing the columns as you go
  - Modified Gram-Schmidt Algorithm
- QR and Singular Matrices
  - Matrix is singular, column of Q is zero.
- Minimization View of QR
  - Basic Minimization approach
  - Orthogonalized Search Directions
  - QR and Length minimization produce identical results
- Mentioned changing the search directions