• Example

\[ X = \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix} \]

\[ X^T X = \begin{pmatrix} 1 + \delta^2 & 1 \\ 1 & 1 + \delta^2 \end{pmatrix} \]
Look at the fitting matrix in more detail

• Suppose we want to solve via least squares
  \[ \mathbf{A}c = y \]
  – \( \mathbf{A} \) is a \( m \times n \) matrix with \( m > n \)

• One way to solve was via LU decomposition of normal equations
  – Poor condition numbers and so not recommended
  – Requires matrix-matrix multiplication which is expensive

• Instead
  – Look for methods that can directly operate on \( \mathbf{A} \) to get the solution
  – Recall in LU we did a set of transformations to \( \mathbf{A} \) and the r.h.s. to find \( \mathbf{c} \)
  – Today we will look at the QR decomposition
Null Space of A

• Look at structure of rectangular systems
• Here $A$ is a matrix that takes $n$ vectors into $m$ vectors with $n < m$
• Not all $m$ vectors will be reachable even if we supply arbitrary $n$ vectors because $A$ is a linear transform
  – $Range$ of $A$: the part of the space of $m$ vectors that are reachable
    $$Range(A) = \{ y \in R^m : y = Ax \text{ for some } x \in R^n \}$$
  – The range of $A$ contains all those vectors that can be made up using the columns of $A$
  – $Rank(A)$ is the dimension of the range of $A$
  – Null space of $A$: those vectors $x$, for which $Ax$ is zero
    $$Null(A) = \{ x \in R^n : Ax = 0 \}$$

$$\text{Dim}(Null(A)) + \text{Rank}(A) = n$$
Null Space of $A^t$

- $A^t$ is a matrix that takes $m$ vectors into $n$ vectors
- Not all $n$ vectors may be reachable even if we supply arbitrary $m$ vectors
  - $Range$ of $A^t$: the part of the space of $n$ vectors that are reachable
    \[ \text{Range}(A^t) = \{ y \in \mathbb{R}^n : y = A^t x \text{ for some } x \in \mathbb{R}^m \} \]
  - The range of $A^t$ contains all those vectors that can be made up using the rows of $A$
  - $Rank(A^t)$ is the dimension of the range of $A^t$
  - Null space of $A^t$: those vectors $x$, for which $A^t x$ is zero
    \[ \text{Null}(A^t) = \{ x \in \mathbb{R}^m : A x = 0 \} \]

\[ \text{Dim(Null}(A^t)) + \text{Rank}(A^t) = m \]
Orthogonal Matrices

- Orthogonal matrices are square matrices that have their columns orthonormal to each other
  - dot product of different column vectors is zero, while of the same column is one
  - Denoted \( Q \)
  - Most trivial orthogonal matrix is the identity matrix
  - \( Q^tQ=I \)

So \( Q^{-1}=Q^T \)

generalization: a complex matrix is *Hermitian* iff \( Q^{-1}=Q^H \)
where superscript \(^H\) denotes complex conjugate transpose
QR decomposition

• Suppose we can write

\[ A = Q'R' \]

- \( Q' \) is an orthonormal matrix of dimension \( m \times m \)
- \( R' \) is a \( m \times n \) matrix that can be written as

\[
\begin{bmatrix}
R \\
0
\end{bmatrix}
\]

\( R \) is a triangular \( n \times n \) matrix and \( 0 \) is a matrix of zeroes of size \( m-n \times n \)

\( Q' \) can also be partitioned as \( [Q \ Q^\sim] \) with \( Q \) containing \( n \) orthonormal columns of size \( m \) and \( Q^\sim \) \( m-n \) orthonormal columns

• If \( Ax = b \) then \( (Q' \ R')x = b \) or \( Q'(R'x) = b \) or \( Q'y = b \)
  - So if \( b \) is in range(\( A \)), it is also in range(\( Q' \))
  - Similarly if \( Q'y = b \); then \( b = Ax \) with \( x = R^{-1}y \)
  - Columns of \( Q \) form an orthonormal basis for range(\( A \))
Orthogonal matrix facts

• Suppose $Q$ is an orthonormal matrix
• Then for any vector $r$ the Euclidean norm is preserved in an orthonormal transformation

Proof

\[ \|Qr\|^2 = (Qr)^t (Qr) = r^t Q^t Q r = r^t (Q^t Q) r = r^t r = \|r\|^2 \]

• If $Q$ is an orthonormal matrix so is the extended matrix $Q_e$

\[ Q_e = \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix} \]

• Easy to show from definition that

\[ Q_e^t Q_e = I \]
$Q^\sim$ forms Nullspace of $(A^t)$

- Choose $z$ in nullspace of $A^t$
- Let $A^t z = 0$
  - $(Q'R')^t z = R^t Q^t z = 0$
  - So $R^t y = 0$ for $y = Q^t z$
  - If $R$ is full rank this means $y$ has to be the zero vector
  - So $Q^t z = 0$
  - So $z$ must be composed of the elements from $Q^\sim$
  - So the columns of $Q^\sim$ form an orthonormal basis for $\text{nullspace}(A^t)$
Solving least squares with QR

- \( A = Q'R' \)
- Let \( r = b - Ax \) \( c = Q'^t b \)
- Goal of least squares find the \( x \) that minimizes squared error (residue)
- Partition \( c \) in to two pieces
  - \( c_1 \) of dimension \( n \)
  - \( c_2 \) of dimension \( m-n \)
- \( \|r\|^2 = \|b - Ax\|^2 = \|b - Q'R'x\|^2 \)
- Length is not changed by multiplication with orthogonal matrix
- So \( \|r\|^2 = \|Q'^t r\|^2 = \|Q'^t [b - Q'R'x]\|^2 \)
  - \( = \|c_1 - Rx\|^2 + \|c_2 - 0x\|^2 \)
- So no matter what \( x \) is the second term remains unchanged
- If we minimize \( \|r\|^2 \) the best we can do is minimize first term
Solving LS via QR

• How do we minimize $\|c_1 - Rx\|^2$
  – If $R$ is full rank set solve $Rx = c$ then we have done the best we can
  – (if $R$ is rank deficient solve in least squares sense)
  – Recall $R$ is triangular so this equation can be easily solved

• Algorithm
  – Compute QR factorization of $A = Q'R'$
  – Form $c_1 = Q^t b$
  – Solve $Rx = c_1$
  – If $R$ is full rank and $Q^t$ is available then the norm of the residual is $\|Q^t b\|$. Else $r = b - Ax$. 