

Look at the fitting matrix in more detail

- Suppose we want to solve via least squares

$$Ax=b$$

- 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
- Today's class
 - Look for methods that can directly operate on A to get the solution
 - Recall in LU we did a set of transformations to A and the r.h.s. to find x
 - Today we will look at the QR algorithm

Null Space

- Today: Other matrix decompositions that are more stable and less expensive
- 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
$$\text{Range}(A) = \{y \in R^m : y = Ax \text{ for some } x \in R^n\}$$
 - 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
$$\text{Null}(A) = \{x \in R^n : Ax = 0\}$$

$$\text{Dim}(\text{Null}(A)) + \text{Rank}(A) = n$$

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
 - $Q^t Q = I$

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 orthogonal 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 orthogonal columns and Q^\sim $m-n$ orthogonal columns

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

Orthogonal matrix facts

- Suppose Q is an orthogonal matrix
- Then for any vector r the Euclidean norm is preserved in an orthogonal 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 orthogonal matrix so is the extended matrix Q_e
- Easy to show from definition that

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

$$Q_e^t Q_e = I$$

Q^{\sim} forms Nullspace of (A^t)

- Choose z in nullspace of A^t
- Let $A^t z = 0$
 - $(Q^t R^t)^t z = R^t Q^t z = 0$
 - 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 $\|r\|^2 = \|Q'^t r\|^2 = \|Q'^t [b - Q' R' x]\|^2 = \|c - \begin{bmatrix} R \\ 0 \end{bmatrix} x\|^2$
 $= \|c_1 - R x\|^2 + \|c_2 - 0x\|^2$

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 $\|c_1 - R x\|^2$
 - 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
 - Form $c_1=Q^t b$
 - Solve $Rx=c_1$
 - If R is full rank and Q^{\sim} is available then the norm of the residual is $\|Q^{\sim t} b\|$. Else $r = b - A x$.

Computing the factorization

- QR is useful ... so how do we factorize a matrix A ?
- In LU we computed an upper triangular matrix by computing adding multiples of other rows so that elements below a given column were zeroed out
- The multipliers were stored in L which gave us $A=LU$
- Here we want to zero out entries below the diagonal but do it with orthogonal matrices
- Two strategies
- Zero out a column at a time using a matrix Q_1 so that $Q_1^t A$ gives us all entries below a certain one in a column as zero
 - Householder transformations
 - Result $Q_n^t \dots Q_2^t Q_1^t A = R$ or $A = Q_1 \dots Q_{n-1} Q_n R = Q R$
- Zero out one specific entry of a column at a time
 - Givens rotations
- Product of orthogonal matrices is orthogonal