Math 381 - Fall 2022

Jay Newby

University of Alberta

Week 9

Last Week

1. Direct solvers for linear systems
2. Stability and backward stability
3. Error bounds
4. Backward/forward substitution
5. LU decomposition
6. Gaussian Elimination (tridiagonal matrices)
7. Gaussian Elimination (general case)
8. Gaussian Elimination with partial pivoting

This Week

1. Linear least squares problems
2. Matrix subspaces (i.e., the range, nullspace, and orthogonal complements)
3. The pseudo inverse of a matrix
4. The QR decomposition
5. Gram-Schmidt orthogonalization

Permutation Matrix
$PP^T = I$

Daily Linear Algebra

Projection Matrices

A matrix $P \in \mathbb{R}^{n\times n}$ is called a *projection matrix* if $P = P^2$.

Complementary Projector

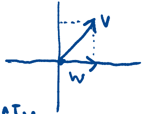
If P is a projection matrix then $(I - P)$ is called its *Complementary Projector*.

Orthogonal Projector (not to be confused with orthogonal matrix!)

Let P be a projection matrix. If P is also symmetric, so that $P = P^T$, then it is an *Orthogonal Projector*, with the property that $y = Px$ and $z = (I - P)x$ implies that $y^T z = 0$ for any $x \in \mathbb{R}^n$.

Example:

$\|\hat{u}\| = 1$ $\quad M = y\hat{u}^T \neq M^T = \hat{u}y^T$



$z = [I - \hat{u}\hat{u}^T]v$

$w = (v^T\hat{u})\hat{u} = (\hat{u}\hat{u}^T)v$

$z^T w = v^T[I - \hat{u}\hat{u}^T]\hat{u}\hat{u}^T v$

$= v^T\hat{u}\hat{u}^T v - v^T(\hat{u}\hat{u}^T)^2 v$

$= v^T\hat{u}\hat{u}^T v - v^T\hat{u}\hat{u}^T v = 0$

$A = \hat{u}\hat{u}^T = A^T = \hat{u}\hat{u}^T$

Daily Linear Algebra

Constructing orthogonal projectors

Any orthonormal set of vectors forming a matrix U can form a orthogonal projection matrix with

$$P = UU^T.$$

Motivation: Linear Least Squares problems

Linear least squares problem

Let $A \in \mathbb{R}^{m\times n}$ be a singular matrix. We wish to find the vector $x \in \mathbb{R}^n$ which yields

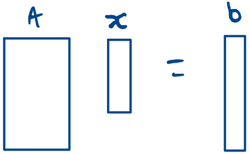
Full rank

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2.$$

residual

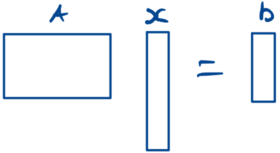
$r = Ax - b$

Overdetermined $(m>n)$



Can have exact sol^n

Underdetermined $(m<n)$

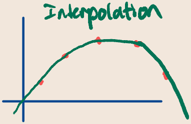


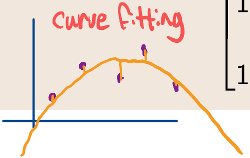
cannot be full rank

Motivation: Linear Least Squares problems

Example: curve fitting (related to interpolation)

Suppose we want to fit a polynomial of degree $n-1$ to the m data points x_i, y_i, where it is assumed that $n < m$. This can be formulated as a least squares solution to





$$\begin{bmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ 1 & x_3 & \cdots & x_3^{n-1} \\ & & & \\ 1 & x_m & \cdots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \approx \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}$$

Vandermond Matrix

Review: matrix subspaces

Range of a matrix

Let $A \in \mathbb{R}^{m \times n}$. The range of A is the space spanned by the column vectors of A. In other words, the vector $Ax \in R(A)$ for all $x \in \mathbb{R}^n$.

Nullspace of a matrix

The nullspace of a matrix $A \in \mathbb{R}^{m \times n}$ contains all null vectors. In other words, if $A\rho = 0$ then $\rho \in N(A)$.

$$\dim(R(A)) + \dim(N(A)) = n$$

Full rank: $\dim(N(A)) = 0$

$\dim(R(A)) = n$

$$A = U\Sigma V^T \Rightarrow AV = U\Sigma$$

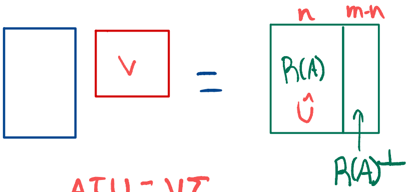
Singular value decomposition (SVD)

For every matrix $A \in \mathbb{R}^{m\times n}$ there exists unitary matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{m\times n}$ such that $A = U\Sigma V^T$.

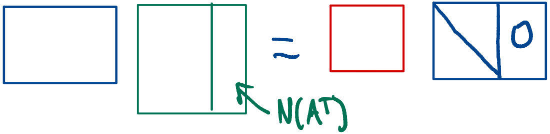
Full Rank case

$$AV = U\Sigma$$

$$N(A^T) = R(A)^\perp$$



$$A^TU = V\Sigma$$



Theorem: Fredholm Alternative

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The equation $Ax = b$ has a solution (not necessarily unique) if and only if $\eta^T b = 0$ for every $\eta \in N(A^T)$.

$b \in R(A)$

Corollary

If a solution to $Ax = b$ exists, it is either the only solution or there are infinitely many solutions.

Solution to the least squares problem

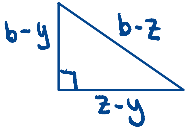
Use projections to find the minimizer

The equation $Ax = Pb$ has a solution if P is a projection onto $R(A)$. The length of residual $r = Ax - b$ is minimized if P is an *orthogonal* projection onto $R(A)$ so that $r = (I - P)b \in R(A)^\perp$.

Proof: Let $y = Pb \in R(A)$. Suppose there is a vector $z \in R(A)$ $\therefore$ $z \neq y$. Want to show $\|b-z\|_2^2 > \|b-y\|_2^2$.

* $[I-P]b = b - y \in R(A)^\perp$

* $z - y \in R(A)$ $(z-y \neq 0)$



$\|b-z\|_2^2 = \|b-y\|_2^2 + \|z-y\|_2^2 > \|b-y\|_2^2$

unless $\|z-y\|_2 = 0$

Solution to the least squares problem

Definition: Pseudo Inverse

For a given full rank matrix $A \in \mathbb{R}^{m\times n}$, its pseudo inverse A^+ is a matrix such that $x = A^+b$ is the unique solution to the least squares problem

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2.$$

The pseudo inverse can be written in many equivalent ways

Normal equations

Normal equations

For a given full rank matrix $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$, the least squares solution to $Ax = b$ is the solution to

$$A^T Ax = A^T b.$$

Proof: $r = Ax - b$, where r is the residual.

$$A^T r = A^T(Ax-b) = A^TAx - A^Tb = 0$$

$\Rightarrow$ either $r=0$ so that $Ax=b$ or $r \in N(A^T) = R(A)^\perp$. By slide 11, x is soln to least squares problem.

Normal equations

Normal equations version of the pseudo Inverse

For a given full rank matrix $A \in \mathbb{R}^{m \times n}$, its pseudo inverse is given by

$$A^+ = (A^T A)^{-1} A^T.$$

Proof: If $A^T A$ is nonsingular (see next slide), and if AA^+ is an orthogonal projection for $R(A)$ (see two slides down) then

$$Ax = AA^+ b \qquad Ax = y = Pb$$

has a solution that solves the LS problem. Then we have that

$$M = A^T A \in \mathbb{R} \qquad Mx = A^T A x = A^T A A^+ b$$

has a unique solution given by

$$x = (A^T A)^{-1} (A^T A) A^+ b = A^+ b.$$

Normal equations

Claim:

If $A \in \mathbb{R}^{m\times n}$ is full rank then A^TA is nonsingular

Proof: Want to show $A^TAp = 0 \Rightarrow p = 0$.

Since is full rank $y = Ap \neq 0$ if $p \neq 0$.

Then we want to show $A^Ty \neq 0$ if $y \neq 0$.

Since $y \in R(A)$ & $R(A)^{\perp} = N(A^T) \Rightarrow A^Ty = 0$ iff $y = 0$.

Normal equations

For a given full rank matrix $A \in \mathbb{R}^{m\times n}$, an orthogonal projector onto $R(A)$ is given by

$$P = AA^+.$$

1. $P^2 = P$ $\quad A\underbrace{\cancel{(A^TA)^{-1}A^TA}}_{I}(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = AA^+$
2. $P^T = P$ $\quad (A(A^TA)^{-1}A^T)^T = A[(A^TA)^{-1}]^TA^T = A(A^TA)^{-1}A^T$
3. $Px \in R(A)$ for any $x \in \mathbb{R}^n$

$$A(A^+x) \in R(A)$$

General approach

Given an orthogonal projector P onto $R(A)$:

1. Compute $y = Pb$
2. Solve $Ax = y$