Lecture 12 Outline

Wednesday, October 2, 2019 12:52 PM

LU Decomposition

Projections

Least Square Approximations

Midterm Exam

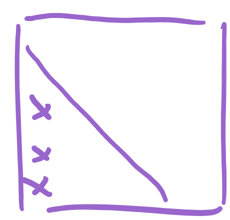
W, Oct 16 in class

LU Pseudocode

Let $\underline{A} \in M_{mn}$ Note: Not necessarily a square matrix

$L = I_{mp}$ with $p = \min(m, n)$

{ 1's on diagonals
{ 0's elsewhere



```
U = A
for i = 1 : min(m-1, n)
    for j = i+1 : m
        L(j,i) = U(j,i)/U(i,i)        1 op
        U(j,i) = 0
        for k = i+1 : n
            U(j,k) = U(j,k) - L(j,i) * U(i,k)
        end
    end
end
```

pivot (at $U(i,i)$)

2ops (for $U(j,k) = U(j,k) - L(j,i) * U(i,k)$)

```
if m>n
  U = U(1:n,1:n)
end
```

Next, consider total operation counts in terms of floating point operations (FLOPS)

Let $m=n$ (square matrix)

T: Total operation count

Focus on loops in pseudocode

$$T=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(1+\sum_{k=i+1}^{n}2\right)$$

$$T=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}1+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sum_{k=i+1}^{n}2$$

$$T=\sum_{i=1}^{n-1}(n-i)+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}2(n-i)$$

$$T=\sum_{i=1}^{n-1}(n-i)+\sum_{i=1}^{n-1}2(n-i)(n-i)$$

$$T = \sum_{i=1}^{n-1} \left[(n-i) + 2(n-i)(n-i) \right]$$

$$T = \sum_{i=1}^{n-1} \left[2n^2 - 4ni + n + 2i^2 - i \right]$$

However,

$$\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$$

$$\sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}$$

$$\therefore T = 2n^2(n-1) - 4\frac{n(n-1)}{2}n + n(n-1) + \frac{2n(n-1)(2n-1)}{6} - \frac{(n-1)n}{2}$$

$$T = \frac{2n^3}{3} - \frac{n^2}{2} - \frac{n}{6}$$

Introduce "Big O" notation: A function

$f(x)$ is $\mathcal{O}(g(x))$, if as $x \to a$ there exists δ and M, such that

$$|f(x)| \leq M\,|g(x)| \text{ for } |x-a| < \delta$$

For the operation count, this means that as n becomes large, the operation count T becomes dominated by the n^3 term

$$\Rightarrow T = \frac{2n^3}{3} + \mathcal{O}(n^2)$$

Thus, in performing LU decomposition, the computational cost scales as n^3

However, once you have $\underline{A} = \underline{L}\,\underline{U}$, solving $\underline{A}\,\underline{x} = \underline{b} \Rightarrow \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$ is only $\mathcal{O}(n^2)$

$\therefore$ Factorization is the expensive part of LU

Can anything go wrong?

Failure of Gaussian Elimination

Consider $\underline{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

Full rank $\rightarrow$ $\underline{A}^{-1}$ exists

$\kappa(\underline{A}) \cong 2.62$

Problem 1: How to eliminate the (2,1) location?

Problem 2: Consider $\underline{A}$ with a slight perturbation

Let $\underline{A} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}$

Exact LU decomposition is

$\underline{L} = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}$, $\underline{U} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{bmatrix}$

$$\underline{L} = \begin{bmatrix} 10^{20} & 1 \end{bmatrix}, \quad \underline{U} = \begin{bmatrix} 0 & 1-10^{20} \end{bmatrix}$$

However, $1-10^{20}$ cannot be represented exactly in finite (double) precision, i.e.

$$1-10^{20} \sim -10^{20}$$

Then, approximate LU

$$\underline{\tilde{L}} = \underline{L}, \quad \underline{\tilde{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & -10^{20} \end{bmatrix}$$

$$\underline{\tilde{L}}\,\underline{\tilde{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 0 \end{bmatrix} \neq \underline{A}$$

Now what?

Pivots to the rescue!

A pivot matrix is one that simply swaps rows. (Technically, this is partial pivoting)

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \underline{U}$$

$$\underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{\underline{P}} \underbrace{\begin{bmatrix} 1 & 1 \end{bmatrix}}_{\underline{A}} = \begin{bmatrix} 0 & 1 \end{bmatrix} = \underline{U}$$

When partial pivoting is used with LU, then one actually has

$$\underline{P}\,\underline{A} = \underline{L}\,\underline{U}$$

Solving $\underline{A}\,\underline{x} = \underline{b}$

$$\underline{P}\,\underline{A}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\underline{L}\,\underline{U}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\Rightarrow \underline{x} = \underline{U}^{-1}\underline{L}^{-1}(\underline{P}\,\underline{b})$$

Check perturbed case:

$$\underline{A} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}, \quad \underline{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\underline{P}\,\underline{A} = \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$$

Let $\underline{E}_1 = \begin{bmatrix} 1 & 0 \\ -10^{-20} & 1 \end{bmatrix}$

$$\rightarrow \underline{E}_1^{-1} = \begin{bmatrix} 1 & 0 \\ 10^{-20} & 1 \end{bmatrix} = \underline{L}$$

$$\underline{E}_1\,\underline{P}\,\underline{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1-10^{-20} \end{bmatrix} = \underline{U}$$

$$\underline{\tilde{U}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\underline{P}\,\underline{A} \simeq \underline{L}\,\underline{\tilde{U}} = \begin{bmatrix} 1 & 0 \\ 10^{-20} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1+10^{-20} \end{bmatrix} \simeq \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$$

Then $\underline{P}^{-1}\,\widetilde{\underline{P}\,\underline{A}}$ … $\begin{bmatrix} 0 & 1 \\ \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 10^{-20} & \end{bmatrix}$

Then $\underline{P}^{-1}\widetilde{\underline{P}\,\underline{A}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$

$$= \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix} = \underline{A} \checkmark$$

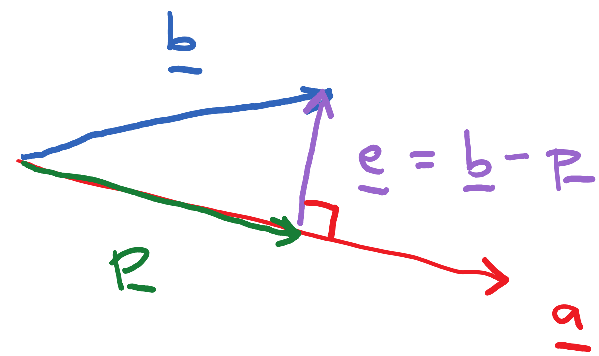
Perturbed matrix!

Projections

Sunday, October 6, 2019 11:34 AM

First consider projections onto a vector (line), then generalize onto a subspace.

A vector projection is the determination of which part of one vector lies on another vector



$\underline{p} + \underline{e} = \underline{b}$

$\underline{e} = \underline{b} - \underline{p}$

$\underline{b}$: Generic vector in $\mathbb{R}^n$

$\underline{a}$: Another vector in $\mathbb{R}^n$

$\underline{p}$: Projection of $\underline{b}$ onto $\underline{a}$

$\underline{e}$: "Error" vector indicating how far $\underline{b}$ is from $\underline{a}$ $\sim$ $\underline{e} \perp \underline{a}$

To find $\mathbf{p}$, first define it as

$$\mathbf{p} = \hat{x}\,\mathbf{a}$$

where $\hat{x}$ is the relative distance along $\mathbf{a}$

Then, the error

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \mathbf{b} - \hat{x}\,\mathbf{a}$$

but

$$\mathbf{e} \perp \mathbf{a} \Rightarrow \mathbf{a} \cdot \mathbf{e} = 0 \Rightarrow \mathbf{a}^T \mathbf{e} = 0$$

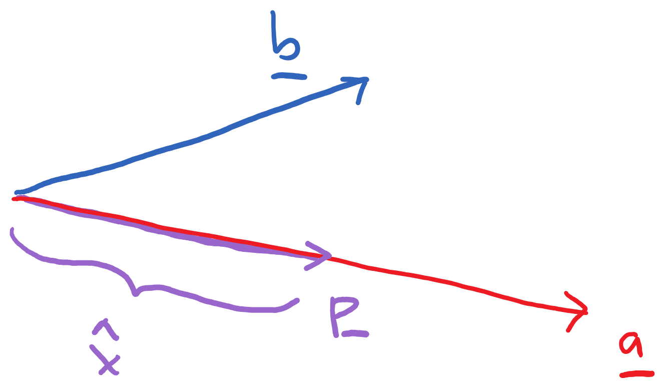
$$\mathbf{a}^T \mathbf{e} = \mathbf{a}^T (\mathbf{b} - \mathbf{p}) = \mathbf{a}^T (\mathbf{b} - \hat{x}\,\mathbf{a}) = 0$$

$$= \mathbf{a}^T \mathbf{b} - \underbrace{\hat{x}}_{\text{scalar}}\, \mathbf{a}^T \mathbf{a} = 0$$

$$\Rightarrow \hat{x} = \frac{\mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}}$$

$$\therefore \mathbf{p} = \hat{x}\,\mathbf{a} = \left(\frac{\mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}} \right) \mathbf{a}$$

scalar



Shortcoming: $\underline{p}$ is only for a specific $\underline{b}$

How can this be generalized?

Projection Matrix: A matrix $\underline{A}$ such that any given vector $\underline{b}$ can be projected onto $\underline{a}$ via

$$\underline{p} = \underline{A}\,\underline{b}$$

$$\underline{p} = \underbrace{\frac{\underline{a}^T \underline{b}}{\underline{a}^T \underline{a}}}_{\text{scalar}} \underline{a} = \underline{a} \frac{\underline{a}^T \underline{b}}{\underline{a}^T \underline{a}} = \underbrace{\frac{\underline{a}\,\underline{a}^T}{\underline{a}^T \underline{a}}}_{\underline{A}\ \text{(matrix)}} \underline{b}$$

Thus, the projection matrix for a vector $\underline{a}$ is given by

$$\underline{A} = \frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}}$$

$\underline{a}^T\underline{a}$: inner product (scalar)

$\underline{a}\,\underline{a}^T$: outer product (matrix)

$\underline{a} \otimes \underline{a}$

Example: In 2D, $\underline{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$

$$\underline{a}\,\underline{a}^T = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} a_1^2 & a_1 a_2 \\ a_1 a_2 & a_2^2 \end{bmatrix}$$

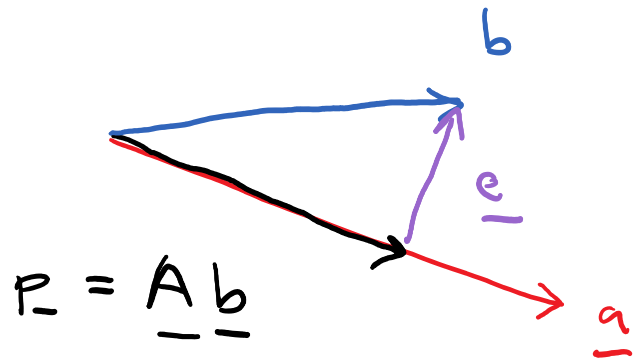
$$\underline{a}^T\underline{a} = a_1^2 + a_2^2$$

$$\underline{A} = \frac{1}{a_1^2 + a_2^2} \begin{bmatrix} a_1^2 & a_1 a_2 \\ a_1 a_2 & a_2^2 \end{bmatrix}$$

Note: $\underline{A}$ will project any vector $\underline{b}$ onto $\underline{a}$

Also, $\underline{A}$ is idempotent: repeated application of $\underline{A}$ has no effect

Example: $\underline{A}(\underline{A}\,\underline{b}) = \underline{A}\,\underline{b}$



$$\underline{A}^2 = \underline{A}\,\underline{A} = \left(\frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}}\right)\left(\frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}}\right) = \frac{\underline{a}\cancel{(\underline{a}^T\underline{a})}\underline{a}^T}{\cancel{(\underline{a}^T\underline{a})}(\underline{a}^T\underline{a})}$$

$$= \frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}} = \underline{A}$$

Furthermore, $\underline{I} - \underline{A}$ can now project onto the perpendicular space of $\underline{a}$

$$(\underline{I} - \underline{A})\underline{b} = \underline{b} - \underline{A}\,\underline{b} = \underline{b} - \underline{p} = \underline{e}$$

Also, $\underline{I} - \underline{A}$ is idempotent

$$(\underline{I} - \underline{A})^2 = (\underline{I} - \underline{A})(\underline{I} - \underline{A})$$

$$= \underline{I}^2 - \underline{I}\underline{A} - \underline{A}\underline{I} + \underline{A}^2$$

$$= \underline{I} - \underline{A} - \underline{A} + \underline{A}$$

$$= \underline{I} - \underline{A}$$

Example: Find the projection matrix of $\underline{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and then project $\underline{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ onto $\underline{a}$

$$\underline{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \underline{A} = \frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}} \qquad \frac{\text{outer product}}{\text{inner product}}$$

$$\underline{a}^T\underline{a} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1^2 + 2^2 = 5$$

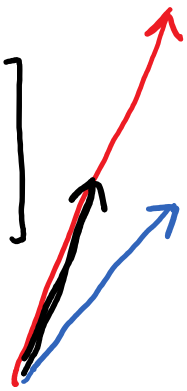
$$\underline{a}\,\underline{a}^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\underline{A} = \frac{\underline{a}\,\underline{a}^T}{\underline{a}^T\underline{a}} = \begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix}$$

$$\underline{p} = \underline{A}\,\underline{b} = \begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix}$$

$$\underline{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\underline{p} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$$



$$\underline{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$