Math 128A, Summer 2019
Lecture 27, Wednesday 8/7/2019

1 Review

We open with going over the first two questions in PSET 6. We have a predictor-corrector scheme. For $k = 1$, we have

$$\begin{aligned} v_{n+1} &= u_n + hf_n, \qquad p(1) = 1 \\ u_{n+1} &= u_n + hf(t_{n+1}, v_{n+1}), \qquad q(1) = 1 \\ [f_{n+1} &= f(t_{n+1}, u_{n+1})], \qquad p = h\left[\tfrac{3}{2} \quad \tfrac{-1}{2}\right] \\ v_{n+1} &= u_n + h\left(\frac{3}{2}f_n - \frac{1}{2}f_{n-1}\right) \\ u_{n+1} &= u_n + h\left(\frac{1}{2}f(v_{n+1} + \frac{1}{2}f_n\right), \qquad q = h\left[\tfrac{1}{2} \quad \tfrac{1}{2}\right] \end{aligned}$$

In `[p,q] = pcoeff(k,t,n)`, we technically can extract n from k, t; however, we don't want to program it this way, because we may have a very long vector, and n helps us know where we are in the vector.

2 Symmetric Positive Definite Matrices

This is a huge topic, and there are probably many books on this subject alone. Recall that the motivation behind this class of matrix is that we **don't need pivoting** for Gaussian elimination.
Suppose we have $x^T Ax > 0$, so that A is symmetric-positive-definite. Inspecting deeper,

$$A = \begin{bmatrix} a_{11} & b^T \\ b & A_{22} \end{bmatrix}$$

where A_{22} is a leading minor submatrix and must be also symmetric-positive definite.
From one step of Gaussian elimination,

$$M_1 A = \begin{bmatrix} a_1 1 & b^T \\ 0 & A_{22} - \frac{1}{a11} bb^T \end{bmatrix}$$

We want to show that one step of Gaussian elimination leaves the resulting matrix symmetric and positive definite. The problem here is that the matrix above is not symmetric! So we consider something that is structurally (destined to be) symmetric, $M_1 A M_1^T$.
Take M_1 as (and check):

$$M_1 = \begin{bmatrix} 1 & 0^T \\ -\frac{b}{a_{11}} & I \end{bmatrix}$$

and hence

$$M_1 A = \begin{bmatrix} 1 & 0^T \\ -\frac{b}{a_{11}} & I \end{bmatrix} \begin{bmatrix} a_{11} & b^T \\ b & A_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & b^T \\ 0 & A_{22} - \frac{1}{a_{11}} bb^T \end{bmatrix},$$

which is the result we had before, so our M_1 is correct. Now we evaluate:

$$M_1 A M_1^T = \begin{bmatrix} a_{11} & b^T \\ 0 & A_{22} - \frac{1}{a_{11}} bb^T \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{a_{11}} b^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} a_{11} & 0^T \\ 0 & A_{22} - \frac{1}{a_{11}} bb^T \end{bmatrix},$$

and our question is if this is symmetric positive definite (we first check if it is symmetric).
The reason $bb^T \geq 0$ is that

$$x^T(bb^T)x = (x^Tb)(b^Tx) = (b^Tx)^2 \geq 0.$$

Recall that $A_{22} > 0$ (submatrix of symmetric positive definite). Let $y \in \mathbb{R}^{n-1}$, so that

$$y^T\left(A_{22} - \frac{1}{a_{11}}bb^T\right)y,$$

and augment matrices to have:

$$\underbrace{\begin{bmatrix} 0 & y^T \end{bmatrix}}_{x^T} \underbrace{\begin{bmatrix} a_{11} & 0 \\ 0 & A_{22} - \frac{1}{a_{11}}bb^T \end{bmatrix}}_{M_1AM_1^T} \underbrace{\begin{bmatrix} 0 \\ y \end{bmatrix}}_{x\in\mathbb{R}^n} = x^TM_1AM_1^Tx = \left(M_1^Tx\right)^T A\left(M_1^Tx\right) > 0,$$

particularly if $y \neq 0$.
Now $A > 0$ implies

$$M_1AM_1^T = \begin{bmatrix} a_{11} & 0^T \\ 0 & A_{22} - \frac{1}{a_{11}}bb^T \end{bmatrix} = \begin{bmatrix} a_{11} & 0^T \\ 0 & M_2\begin{bmatrix} a_22' & 0^T \\ 0 & A_{33}' - \frac{1}{a_{11}'}b'b'^T \end{bmatrix}M_2^T \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \\ & M_2 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} & 0^T \\ 0 & \underbrace{A_{33}''}_{>0} \end{bmatrix} \begin{bmatrix} 1 & \\ & M_2^T \end{bmatrix}$$

So we deduce:

$$M_{n-1}\cdots M_1AM_1^T\cdots M_{n-1}^T = \begin{bmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{bmatrix}$$

which we call this
Symmetric unit-lower-triangular reduction to diagonal form, also known as LDL^T factorization or **Cholesky factorization**.
That is, we have

$$A = LDL^T,$$

where D has entries $d_{ii} > 0$. And if $A = A^T$, it's plausible to say we have a theorem:

Theorem 2.1.

$$\begin{aligned} A = A^T > 0 &\iff x^TAx > 0\forall_{x\neq 0} \\ &\iff A = L^TDL, \quad L \text{ is unit lower-triangular, } D > 0 \text{ diagonal} \\ &\iff \exists_{B \text{ invertible}} A = B^TB \\ &\iff \exists_{R \text{ upper-triangular}} A = R^TR \quad (R^TR) \end{aligned}$$

After the break we'll talk about 3 different ways to compute the Cholesky factorization.

Break time.

3 Cholesky Factorization

This is symmetric Gaussian Elimination without Pivoting. Because we know we are not pivoting, we are really just solving a bunch of equations for a symmetric matrix:

$$A = R^T R$$

$$\begin{bmatrix} \ddots & & \\ & a_{ij} & \\ & & \ddots \end{bmatrix} = \begin{bmatrix} r_{11} & & 0 \\ & \ddots & \\ r_{1n} & & r_{nn} \end{bmatrix}$$

Getting these matrices satisfied is just $\frac{n(n_1)}{2}$ quadratic equations. Due to the ('symmetric') form of A and correspondingly the lower-triangular times upper-triangular on the RHS, we have a chance at finding an order in which we can isolate and solve for one at a time. That is,

$$\begin{aligned} a_{11} &= r_{11}^2 \\ a_{12} &= r_{11} r_{12} \\ &\vdots \\ a_{1n} &= r_{11} r_{1n} \end{aligned}$$

which gives us a full row of R and a column of R^T, so we have a lot of information already! Next we have:

$$\begin{aligned} a_{22} &= r_{12}^2 + r_{22}^2 \\ a_{23} &= r_{12} r_{13} + r_{22} \underbrace{r_{23}}_{\text{only unknown}} \\ &\vdots \end{aligned}$$

Example: Our favorite symmetric positive definite matrix.

Actually, we cheat by taking:

$$\begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -2 & 13 \end{bmatrix} > 0$$

and we constructed a symmetric positive definite matrix. Now how do we find its Cholesky factorization?

$$\begin{bmatrix} 1 & -2 \\ -2 & 13 \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{bmatrix},$$

which gives us the equations:

$$\begin{aligned} a^2 = 1 &\implies a = 1 \\ ab = -2 &\implies b = -2 \\ b^2 + c^2 = 13 &\implies c = 3, \end{aligned}$$

which gives us our original matrices.

3.1 Can Finding the Cholesky Factorization Fail?

It turns out, this only fails if A is not Symmetric-Positive-Definite.
If this fails, given R, insert E such that $(R+E)$ satisfies the Cholesky factorization:

$$(R+E)^T(R+E) = A$$

This gives:

$$\begin{aligned} R^TR + R^TE + E^TR + E^TE &= A \\ R^TE + E^TR &= A - R^TR - \underbrace{\cancel{E^TE}}_{O(E^2)} \end{aligned}$$

This is one step of Newton's method, where we are linearizing and throwing away the quadratic term. We know that Newton's converges quadratically when it works.

$$\begin{aligned} R^TER^{-1} + E^T &= AR^{-1} - R^T \\ \underbrace{E}_{up.tri.}\ \underbrace{R^{-1}}_{up.tri.} + \underbrace{R^{-T}}_{low.tri}\ \underbrace{E^T}_{low.tri} &= \underbrace{R^{-T}AR^{-1} - I}_{\text{symmetric}} \end{aligned}$$

Note that E and R^{-1} are each upper-triangular and R^{-T}, E^T are each lower-triangular.

Definition: Causality -

Take a lower-triangular matrix L and a time vector $T = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}$, where

$$\begin{bmatrix} \cdot & & \\ \cdot & \cdot & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \\ \\ \\ \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \\ \\ \end{bmatrix}$$

Hence we say:

$$ER^{-1} = \text{uph}(R^{-T}AR^{-1} - I), \tag{1}$$

where

$$\text{uph}(A)_{ij} := \begin{cases} A_{ij}, & j > i \\ \frac{1}{2}A_{ij}, & j = i \\ 0, & j < i \end{cases},$$

where $j = i$ gives the upper triangular with half on the diagonal. So E solves the following equation:

$$(R+E)^T(R+E) = A + E^TE\ \underbrace{(R+E)^T(R+E) - A}_{\text{residual}} \qquad = E^TE$$

Now we want to express a relationship between the residual and the error. We write:

$$E := \mathrm{uph}(R^T A R^{-1} - I)R$$

from (1) above.

and hence

$$\begin{aligned} E^T E &= \left(\mathrm{uph}(R^T A R^{-1} - I)R\right)^T \left(\mathrm{uph}(R^{-T} A R^{-1} - I)R\right) \\ &= R^T \mathrm{uph}(R^{-T} \underbrace{[A - R^T R]}_{\text{old residual}} R^{-1}) \cdot \mathrm{uph}(R^{-T} \underbrace{[A - R^T R]}_{\text{old residual}} R^{-1})R \\ \frac{E^T E}{A} &= \underbrace{O\left((A - R^T R)^2\right)}_{A^2} \underbrace{R^{-2}}_{A^{-1}}, \end{aligned}$$

and so we say relative residual is squaring.

Lecture ends here.