Secant Method

Q: What about methods with $x_{k+1} = g(x_k, x_{k-1})$?

Recall: $$f'(x) \approx \frac{f(x+h) - f(x)}{h} \quad \text{for } h \text{ small.}$$

Def: The secant method is defined by

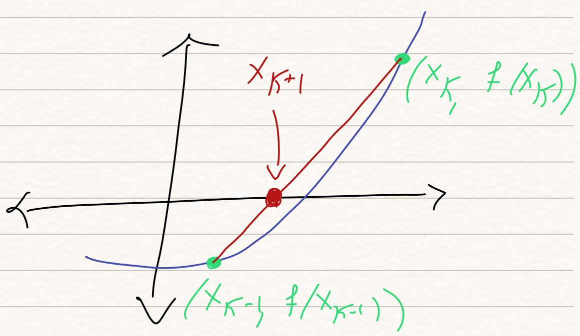
$$x_{k+1} = x_k - f(x_k)\,\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \qquad k = 1, 2, \ldots$$

with x_0, x_1 given starting values.

We assume $f(x_k) \neq f(x_{k-1})$ for any k.

Note: Like Newton, but does not require f'!

Note: See homework for the similar "Steffensen's Method".



Thm: (Convergence of Secant Method)

Say $f \in C^1([\xi - h, \xi + h], \mathbb{R})$ for some $h > 0$.

Say $f(\xi) = 0$, $f'(\xi) \neq 0$.

Then $x_k \to \xi$ with order $q = \frac{1}{2}(1 + 5^{1/2})$ so long as x_0, x_1 are sufficiently close to ξ.

Pf: See book. $\square$

Note: Slower than Newton, but can be cheaper per iteration!

Bisection Method

Let $f \in C^0([a,b], \mathbb{R})$.

Say $\xi \in [a,b]$ with $f(\xi) = 0$. Assume $f(a) f(b) < 0$.

Alg: Set $a_0 = a$, $b_0 = b$.
while $|a_k - b_k| > \varepsilon$:
Compute $c_k = 1/2 (a_k + b_k)$

$$\text{set } (a_{k+1}, b_{k+1}) = \begin{cases} (a_k, c_k) & \text{if } f(c_k) f(b_k) > 0 \\ (c_k, b_k) & \text{if } f(c_k) f(b_k) < 0 \end{cases}$$

Clearly $|c_k - \xi| \le 2^{-k} |c_0 - \xi|$, so $c_k \to \xi$ linearly.
only requires continuity.

Solution of Linear Systems

We now consider a very fundamental problem:

set of nxn matrices w/ real entries

(P) Given $\underline{\underline{A}} \in \mathbb{R}^{n \times n}$, $\underline{b} \in \mathbb{R}^n$, find $\underline{x} \in \mathbb{R}^n$ s.t.

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

set of n-vectors

Notation: We write a_{ij} as the ij^{th} element of $\underline{\underline{A}}$.

$$\underline{\underline{A}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix}, \quad \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Recall: We can write $\underline{\underline{A}}\,\underline{x} = \underline{b}$ as:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}$$

Summation: $$(\underline{\underline{A}}\,\underline{x})_i = \sum_{j=1}^{n} a_{ij}x_j = b_i$$

$= \underline{a}_i \cdot \underline{x}$ w/ $\underline{a}_i = i^{th}$ row of $\underline{\underline{A}}$, i.e.,

$$\underline{\underline{A}} = \begin{pmatrix} - & \underline{a}_1^T & - \\ - & \underline{a}_2^T & - \\ & \vdots & \\ - & \underline{a}_n^T & - \end{pmatrix}$$

Def: We say $\underline{\underline{A}} \in \mathbb{R}^{n\times n}$ is nonsingular if $\det(\underline{\underline{A}}) \neq 0$.

Formally, we have

$$\det(\underline{\underline{A}}) = \sum_{\text{permutations } \nu} \operatorname{sign}(\nu_1, \dots, \nu_n)\, a_{1\nu_1} a_{2\nu_2} \cdots a_{n\nu_n}$$

Recall: Nonsingular matrices have an inverse, $\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}}$

where the identity matrix $\underline{\underline{I}}$ is given by

$$\underline{\underline{I}} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix}, \quad I_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Assuming $\underline{\underline{A}}$ nonsingular, we can "solve" (P) analytically as:

$$\underline{x} = \underline{\underline{A}}^{-1}\underline{b}.$$

Cramer's Rule gives an expression for the components of $\underline{x}$:

$$x_i = \frac{\det(A_i^{\omega})}{\det(A)}$$ matrix with i^{th} column of $\underline{\underline{A}}$ replaced by $\underline{b}$

Requires calculation of determinants.

number of $+/-/\times/\div$ ("flops") scales like $n!$ (factorial)

way too slow for real applications.

Computing $\underline{\underline{A}}^{-1}$ explicitly and multiplying is equally ineffective.

Gaussian Elimination

Ex: $\underline{\underline{A}} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}$, $\underline{b} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}$ $\Rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & 4 & 2 & 16 \\ -1 & 5 & -4 & -3 \end{array}\right)$

first row $\cdot -2$ added to second row:
first row added to third row:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 6 & -3 & 3 \end{array}\right) \xrightarrow{R_3 - 3 \cdot R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -3 & -9 \end{array}\right)$$

Easily solved in reverse order!

$$-3x_3 = -9 \Rightarrow x_3 = 3$$
$$2x_2 = 4 \Rightarrow x_2 = 2$$
$$x_1 + x_3 + x_2 = 6 \Rightarrow x_1 = 1$$

Δ

Claim: These row operations can be expressed via matrix mult.

Let $\underline{\underline{E}}^{(rs)} \in \mathbb{R}^{n \times n}$ be $E^{(rs)}_{ij} = \begin{cases} 1 & i = r \;\; j = s \\ 0 & \text{else} \end{cases}$

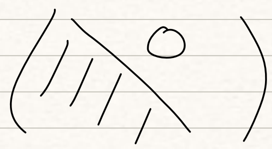
s column s
$1\; 0\; 0 \cdots 0\;\cdot$ row r

$$\underline{\underline{E}}^{(rs)} = \begin{array}{c} \\ r \\ \\ \end{array} \begin{pmatrix} 0 & \cdots & & 0 \\ \vdots & & 1 & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}$$

So $\underline{\underline{E}}^{(rs)} \underline{\underline{A}}$ has all rows $= 0$ except row r, which is the s^{th} row of $\underline{\underline{A}}$!

Then $(\underline{\underline{I}} + \mu^{(rs)} \underline{\underline{E}}^{(rs)}) \underline{\underline{A}}$ adds $\mu^{(rs)} \cdot \underline{a}_s$ to $\underline{a}_r$.

Def: We say $\underline{\underline{L}} \in \mathbb{R}^{n \times n}$ is lower-triangular if $\ell_{ij} = 0$ for all $j > i$. We say that $\underline{\underline{L}}$ is unit lower-triangular if in addition the diagonals $L_{ii} = 1$ for all i.

Picture: $\underline{\underline{L}} =$ 

Note: $\underline{\underline{I}} + \mu^{(rs)} \underline{\underline{E}}^{(rs)}$ is unit lower triangular if $r > s$, which is what we use for Gaussian Elimination (G.E.)

So Gaussian elimination can be expressed by a bunch of left-multiplications of the form $\underline{\underline{I}} + \mu^{(rs)} \underline{\underline{E}}^{(rs)}$.

Q: How many such matrices?

A: # of elements below diagonal $= \frac{1}{2} \cdot n \cdot (n-1)$

Thm: (i) Product of LT matrices is LT

(ii) " " ULT " " ULT

(iii) LT matrices nonsingular if and only if diagonals are all nonzero.

(iv) inverse of LT is also LT

(v) " " ULT " " ULT

Pf: (i–iv) are simple (try at home!). We only prove (v).

We prove by induction on n.

For $n=2$, $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}$

Assume true for all $2 \le n \le k$. Consider $n = k+1$

Let us partition,

$$\underline{\underline{L}} = \begin{pmatrix} \underline{\underline{L}}_1 & \underline{0} \\ \underline{r}^T & \gamma \end{pmatrix} \qquad \underline{\underline{L}}_1 \in \mathbb{R}^{k\times k}, \text{ unit LT}$$

$$\underline{\underline{X}} \in \mathbb{R}^{k\times k}$$

$$\underline{\underline{L}}^{-1} = \begin{pmatrix} \underline{\underline{X}} & \underline{y} \\ \underline{z}^T & \beta \end{pmatrix} \qquad \gamma, \beta \in \mathbb{R}$$

$$\underline{r}, \underline{0}, \underline{y}, \underline{z} \in \mathbb{R}^{k-1}$$

By assumption, $\underline{\underline{L}}\,\underline{\underline{L}}^{-1} = \underline{\underline{I}}_{(k+1)\times(k+1)}$. But,

$$\underline{\underline{L}}\,\underline{\underline{L}}^{-1} = \begin{pmatrix} \underline{\underline{L}}_1 & \underline{0} \\ \underline{r}^T & \gamma \end{pmatrix} \begin{pmatrix} \underline{\underline{X}} & \underline{y} \\ \underline{z}^T & \beta \end{pmatrix} = \begin{pmatrix} \underline{\underline{L}}_1 \underline{\underline{X}} & \underline{\underline{L}}_1 \underline{y} \\ \underline{r}^T \underline{\underline{X}} + \gamma \underline{z}^T & \underline{r}^T \underline{y} + \gamma\beta \end{pmatrix}$$

So we require that:

$$\underline{\underline{L}}_1 \underline{\underline{X}} = \underline{\underline{I}}_{k\times k}, \quad \underline{\underline{L}}_1 \underline{y} = \underline{0}$$

$$\underline{r}^T \underline{\underline{X}} + \gamma \underline{z}^T = \underline{0}^T, \quad \underline{r}^T \underline{y} + \gamma\beta = 1$$

So $\underline{\underline{X}} = \underline{\underline{L}}_1^{-1}$, which is LT of order k by induction hypothesis.

$\underline{\underline{L}}_1$ nonsingular $\Rightarrow$ $\underline{y} = 0$.

Remaining two equations give $\underline{z}, \beta$, noting $\gamma \neq 0$ because $\underline{\underline{L}}$ invertible.

This shows L^{-1} lower triangular of order k $\square$

Def: We say $U \in \mathbb{R}^{n \times n}$ is upper-triangular if $u_{ij} = 0$ for all $i > j$.

Elimination Process

$$L_N L_{N-1} \cdots L_1 A = U, \qquad N = \tfrac{1}{2} \cdot n \cdot (n-1)$$

each L_i LT, U UT!

each $L_i = I + \ell^{(rs)} E^{(rs)}$ for some $r, s, \ell^{(rs)}$.

Note: $E^{(rs)} E^{(rs)} = \delta_{rs} E^{(rs)}$, $\delta_{rs} = \begin{cases} 1 & r = s \\ 0 & \text{else} \end{cases}$

Hence, $\left(I + \ell^{(rs)} E^{(rs)}\right)^{-1} = \left(I - \ell^{(rs)} E^{(rs)}\right)$

Hence,

$$A = \underbrace{\left(L_1^{-1} L_2^{-1} \cdots L_N^{-1}\right)}_{= L} U = LU$$

This is called the LU factorization of A.