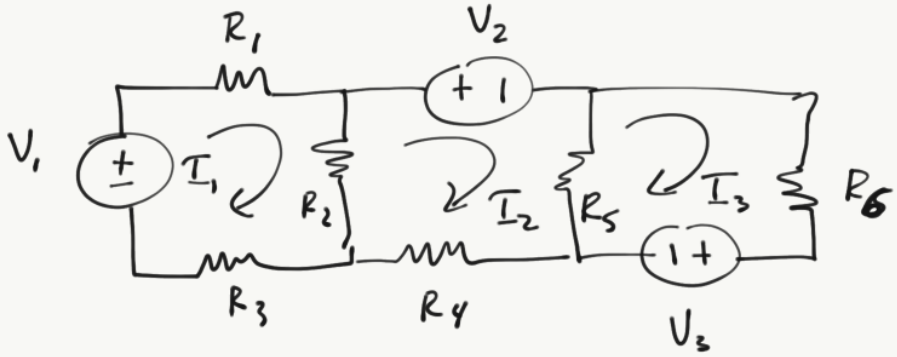
Linear Algebra

- Sets of linear equations play a key role in many physical systems — these can be represented as matrices
- For example analysis of circuits:



(I_i unknown)

Applying Kirchoff's voltage laws around each loop:

$$V_1 - I_1 R_1 - (I_1 - I_2) R_2 - I_1 R_3 = 0$$

$$-V_2 - (I_2 - I_3) R_5 - I_2 R_4 - (I_2 - I_1) R_2 = 0$$

$$-(I_3 - I_2) R_5 - I_3 R_6 - V_3 = 0$$

Rearranging, we find:

$$(-R_1 - R_2 - R_3) I_1 + R_2 I_2 + 0 = -V_1$$

$$+R_2 I_1 + (-R_5 - R_4 - R_2) I_2 + R_5 I_3 = V_2$$

$$+0 + R_5 I_2 + (-R_5 - R_6) I_3 = V_3$$

- More generally a system of 3 eqns. & 3 unknowns:

$$A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = b_1$$

$$A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = b_2$$

$$A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = b_3$$

- These can be represented in matrix form as:

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

(note compactness!)

A_{ij} are elements of $\underline{\underline{A}}$
x_j are elements of $\underline{x}$
b_i are elements of $\underline{b}$

- Referring back to the system above & enforcing equivalence with matrix form, it is clear that matrix multiplication is defined by:

$$\left(\underline{\underline{A}}\,\underline{x}\right)_i \equiv \sum_j A_{ij}x_j$$

→ more generally for two matrices

$$\boxed{\left(\underline{\underline{A}}\,\underline{\underline{B}}\right)_{ij} \equiv \sum_k A_{ik}B_{kj}}$$

matrix multiplication

- For multiplication to make sense so matrices must be conformable, viz. of size: $(m \times p)\cdot(p \times n)$

inner dimensions equal!

$m \times n$

- Mathematically speaking, a system of equations could fall into one of four categories:

  1) no solution (inconsistent)
  2) trivial solution $\underline{x} = \underline{0}$
  3) infinite solutions conforming to some constraint
  4) unique solution

- For physical situations we expect 4) !!!
- Linear systems play a large role in numerical analysis - networks, fitting & estimation of parameters, interpolation, & ODEs/PDEs

- Special Matrices:

  1) column vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}$

  2) row vector $\underline{y} = [y_1 \; y_2 \; \dots \;]$

  by notational default we assume $\underline{x}$ is column & $\underline{x}^T$ is row, unless otherwise stated...

  3) square matrix has size $n \times n$

  4) diagonal matrix $\underline{\underline{D}} = \begin{bmatrix} d_{11} & & 0 \\ & d_{22} & \\ 0 & & \ddots \end{bmatrix}$

identity matrix $\underline{\underline{I}} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \\ & & & 1 \end{bmatrix}$

upper triangular $\underline{\underline{U}} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots \\ & u_{22} & u_{23} & \\ 0 & & \ddots & \end{bmatrix}$

lower triangular $\underline{\underline{L}} = \begin{bmatrix} l_{11} & & 0 \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \cdots \end{bmatrix}$

banded $\underline{\underline{B}} = \begin{bmatrix} b_{11} & 0 & b_{13} & 0 \\ 0 & b_{22} & 0 & b_{24} \\ b_{31} & 0 & b_{33} & 0 \\ 0 & b_{42} & 0 & b_{44} \end{bmatrix}$

- Matrix operations all follow from linear equation properties

$$(\underline{\underline{A}} + \underline{\underline{B}})_{ij} = A_{ij} + B_{ij} \; ; \; (\underline{\underline{A}} - \underline{\underline{B}})_{ij} = A_{ij} - B_{ij}$$

$\underline{\underline{A}}, \underline{\underline{B}}$ must have same size!

$$(\alpha \underline{\underline{A}})_{ij} = \alpha A_{ij}$$

$$\underline{\underline{A}}(\underline{\underline{B}} + \underline{\underline{C}}) = \underline{\underline{A}}\underline{\underline{B}} + \underline{\underline{A}}\underline{\underline{C}} \; ; \; \boxed{\underline{\underline{A}}\underline{\underline{B}} \neq \underline{\underline{B}}\underline{\underline{A}}}$$

- Matrix inverse $\underline{\underline{A}}^{-1}$ defined s.t :

$$\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1}\,\underline{\underline{A}} = \underline{\underline{I}}$$

- Solutions to $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{C}}$ for $\underline{\underline{A}}$:

$$\underline{\underline{A}} = \underline{\underline{B}}^{-1}\underline{\underline{C}}$$

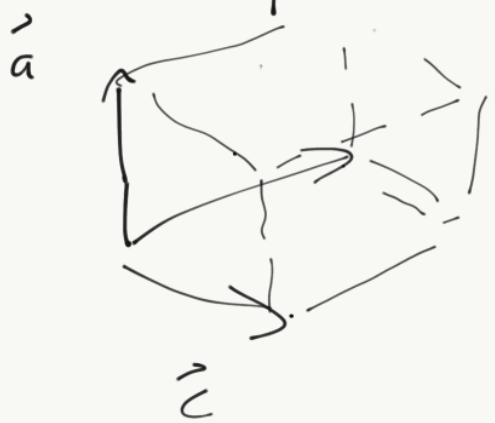
similarly $\underline{\underline{A}}\,\underline{x} = \underline{b} \rightarrow$ $\boxed{\underline{x} = \underline{\underline{A}}^{-1}\,\underline{b}}$

- Determinants of matrices can be thought of as analogous to concept of "volume" or magnitude

$$\underline{\underline{A}} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \qquad \vec{a} = (a_1, a_2, a_3)$$

$\vec{b} = \ldots$ etc.

$\det(\underline{\underline{A}}) \equiv |\underline{\underline{A}}|$ is related to the triple product $\vec{a}\cdot(\vec{b}\times\vec{c})$ which is volume spanned by $\vec{a}, \vec{b}, \vec{c}$



if $|\underline{\underline{A}}| = 0$

then $\vec{a}$ lies in plane of $\vec{b}, \vec{c}$

linearly dependent

- More generally we can define $|\underline{\underline{A}}|$ by:

$$|\underline{\underline{A}}| = \sum_{j=1}^{n} A_{ij}\, \text{cof}(A_{ij}) = \sum_{j=1}^{n} A_{ij}\,(-1)^{(i+j)}\, M_{ij}$$

(i arb) ↙ $j=1$ ↓ cofactor of A_{ij} $j=1$ ↓ minor of A_{ij}

M_{ij} is the minor defined by evaluation det of augmented matrix, without row i, column j

- Note that det(scalar) = scalar.
- Consider a 2×2 determinant

$$|\underline{\underline{A}}| = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}(-1)^{1+1} M_{11} + A_{12}(-1)^{1+2} M_{12}$$

$$M_{11} \rightsquigarrow \begin{vmatrix} \cancel{A_{11}} & \cancel{A_{12}} \\ \cancel{A_{21}} & A_{22} \end{vmatrix} = \det(A_{22}) = A_{22}$$

$$M_{12} \rightsquigarrow \begin{vmatrix} \cancel{A_{11}} & \cancel{A_{12}} \\ A_{21} & \cancel{A_{22}} \end{vmatrix} = \det(A_{21}) = A_{21}$$

$$\boxed{|\underline{\underline{A}}| = A_{11}A_{22} - A_{12}A_{21}} \quad (2\times 2 \text{ det})$$

- Can also expand cofactors along column:

$$\boxed{|\underline{\underline{A}}| = \sum_{i=1}^{n} A_{ij}\, \text{cof}(A_{ij})} \quad (j \text{ fixed})$$

- 3×3 determinants proceed similarly:

$$|\underline{\underline{A}}| = A_{11}(-1)^{1+1}M_{11} + A_{12}(-1)^{1+2}M_{12} + A_{13}(-1)^{1+3}M_{13}$$

$$M_{11} \sim \begin{vmatrix} \cancel{A_{11}} & \cancel{A_{12}} & \cancel{A_{13}} \\ A_{21} & A_{22} & A_{23} \\ \cancel{A_{31}} & A_{32} & A_{33} \end{vmatrix} = A_{22}A_{33} - A_{23}A_{32}$$

$$M_{12} = A_{21}A_{33} - A_{23}A_{31} \;;\; M_{13} = A_{21}A_{32} - A_{22}A_{31}$$

$$|\underline{\underline{A}}| = A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{22}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31})$$

- Evaluating an $n\times n$ determinant requires some adds & $n!$ terms, each containing $(n-1)$ multiplications, so it scales like: adds

$$\mathcal{O}((n-1)n!)$$ (cofactor expansion of det)

Elimination Methods

- Cramer's rule is a direct method & useful for establishing a baseline approach for comparing efficiency of various methods:

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \quad ; \quad x_j = \frac{|\underline{\underline{A}}^j|}{|\underline{\underline{A}}|} \quad ; \quad (\underline{\underline{A}}^j \text{ is } \underline{\underline{A}} \text{ with the } j^{th} \text{ column replaced by } \underline{b})$$

- e.g. $x_1 = \dfrac{\begin{vmatrix} b_1 & A_{12} \\ b_2 & A_{22} \end{vmatrix}}{A_{11}A_{22} - A_{12}A_{21}}$; $\dfrac{\begin{vmatrix} A_{11} & b_1 \\ A_{21} & b_2 \end{vmatrix}}{A_{11}A_{22} - A_{12}A_{21}} = x_2$

$$x_1 = \frac{b_1 A_{22} - A_{12} b_2}{A_{11}A_{22} - A_{12}A_{21}} \; ; \quad x_2 = \frac{A_{11} b_2 - b_1 A_{21}}{A_{11}A_{22} - A_{12}A_{21}}$$

- Dealing with larger systems is problematic. For $\underline{\underline{A}} \sim n \times n$ we must evaluate $n+1$ determinants.
- Each determinant has $n!$ products (all containing n items multiplied) So $(n!)(n-1)$ multiplications per det.
- Total multiplications:

$$|\underline{\underline{M}}| = \mathcal{O}((n+1)(n-1)\,n!)$$

$$= \mathcal{O}((n-1)(n+1)!)$$

- Thus, Cramer's rule is too computationally expensive to practically use. (SHOW IN MATLAB).
- Elimination methods address this shortcoming; these operate very much like a normal solution method by hand.
  1) perform algebraic substitutions until value of one variable can be found (fwd elimination)
  2) substitute this variable back into simplified eqns successively to find other vars. (back substitution)

- Elementary row operations (allow elimination steps)
  1) Any row can be multiplied by scalar (scaling)
  2) Rows can be interchanged (pivoting)
  3) Rows may be linearly combined with other rows. (elimination)
- None of these rows operations change the solution to the system $\underline{\underline{A}}\,\underline{x} = \underline{b}$

- Connection of row operations to matrix multiplication, e.g.:

1) $\underline{\underline{E}}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\underline{\underline{E}}_1 \underline{\underline{A}} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ \alpha A_{21} & \alpha A_{22} & \alpha A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$

2) $\underline{\underline{E}}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\underline{\underline{E}}_2 \underline{\underline{A}} = \begin{bmatrix} A_{21} & A_{22} & A_{23} \\ A_{11} & A_{12} & A_{13} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$

$$\underline{\underline{E}}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{bmatrix}; \quad \underline{\underline{E}}_3 \underline{\underline{A}} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ \alpha A_{11} + A_{31} & \alpha A_{12} + A_{32} & \alpha A_{13} + A_{33} \end{bmatrix}$$

- Because of the connection of row ops to multiplication we can conceive as an inverse as a sequence of row ops.

$$\left.\begin{aligned} \underline{\underline{A}}^{-1} \underline{\underline{A}} &= \underline{\underline{I}} \\ \left(\prod_i \underline{\underline{E}}_i\right) \underline{\underline{A}} &= \underline{\underline{I}} \end{aligned}\right\}$$

so its possible to find an inverse hence solution from row operations!!!

- So if we find row ops that reduce $\underline{\underline{A}}$ to $\underline{\underline{I}}$ we can apply those ops to $\underline{\underline{I}}$ to get $\underline{\underline{A}}^{-1}$:

$$\underline{\underline{A}}^{-1} = \left(\prod_i \underline{\underline{E}}_i\right) \cdot \left(\prod_i \underline{\underline{E}}_i\right) \underline{\underline{I}}$$

NOTE: Matrix multiplication is associative!

$$\underline{\underline{A}}\left(\underline{\underline{B}}\underline{\underline{C}}\right) = \left(\underline{\underline{A}}\underline{\underline{B}}\right)\underline{\underline{C}}$$

$$\begin{aligned} x_1 + 4x_2 + 2x_3 &= 15 \\ 3x_1 + 2x_2 + x_3 &= 10 \\ 2x_1 + x_2 + 3x_3 &= 13 \end{aligned} \rightarrow \left[\begin{array}{ccc|c} 1 & 4 & 2 & 15 \\ 3 & 2 & 1 & 10 \\ 2 & 1 & 3 & 13 \end{array}\right]$$

elim. from row 1:

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 15 \\ 3 & 2 & 1 & 10 \\ 2 & 1 & 3 & 13 \end{array}\right] \quad \begin{array}{l} R_2 - 3R_1 \\ R_3 - 2R_1 \end{array}$$

elim. from row R_2:

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 15 \\ 0 & -10 & -5 & -35 \\ 0 & -7 & -1 & -17 \end{array}\right] \quad R_3 - \tfrac{7}{10} R_2$$

$$\begin{aligned} -1 - \tfrac{7}{10}(-5) &= 5/2 \\ -17 - \tfrac{7}{10}(-35) &= -17 + \tfrac{245}{10} \\ &= \tfrac{-170+245}{10} \\ &= \tfrac{75}{10} = \tfrac{15}{2} \end{aligned}$$

final state following elim:

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 15 \\ 0 & -10 & -5 & -35 \\ 0 & 0 & 5/2 & 15/2 \end{array}\right]$$

backsubstitution:

$$x_3 = (15/2)\left(\tfrac{2}{5}\right) = 3$$

$$x_2 = (-35 + 5(3))\left(\tfrac{-1}{10}\right) = -20\left(\tfrac{-1}{10}\right) = 2$$

$$x_1 = (15 - 2(3) - 4(2)) = 1$$

- Two modifications to standard elimination schemes are advisable: pivoting & scaling.

- The pivot element for an elimination step is the element being used for elimination. e.g.

pivot

$$\left[\begin{array}{ccc|c} \textcircled{3} & 2 & 1 & 10 \\ 2 & 1 & 3 & 13 \\ 1 & 4 & 2 & 15 \end{array}\right] \quad \begin{array}{l} R_2 - \left(\frac{2}{3}\right) R_3 \\ R_3 - \left(\frac{1}{3}\right) R_3 \end{array}$$

elimination multiplier

- Interchanging rows (reordering eqns.) doesn't change solution — so one can switch pivots with impunity

- For reasons of numerical accuracy, it is always best to scale each row by largest element in that row & then switch to largest pivot.

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 15 \\ 3 & 2 & 1 & 10 \\ 2 & 1 & 3 & 13 \end{array}\right] \quad \begin{array}{l} \text{pivot: } \frac{1}{15} \\ \text{pivot: } \frac{3}{10} \\ \text{pivot: } \frac{2}{13} \end{array} \quad \text{swap}$$

$$\downarrow$$

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 10 \\ 1 & 4 & 2 & 15 \\ 2 & 1 & 3 & 13 \end{array}\right]$$

do elimination, then reorder again for largest scaled pivot

defines Gaussian elimination

- Gauss – Jordan elimination is based on the fact that one may compute inverse as a sequence of row ops on $\underline{\underline{I}}$:

$$\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{I}} \rightarrow \left(\prod_i \underline{\underline{E}}_i\right)\underline{\underline{A}} = \underline{\underline{I}}$$

$$\therefore \left(\prod_i \underline{\underline{E}}_i\right)\underline{\underline{I}} = \underline{\underline{A}}^{-1}$$

$$\left[\begin{array}{ccc|ccc} 1 & 4 & 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \end{array}\right] = \left[\underline{\underline{A}} \,\middle|\, \underline{\underline{I}}\right] \quad \begin{array}{l} R_2 - 3R_1 \\ R_3 - 2R_1 \end{array}$$

- See book examples & homework.
- Generally elimination methods of all types are more efficient than solving a system via Cramer's Rule; which scaled like

$$\mathcal{O}((n-1)(n+1)!) \qquad \underline{\text{Cramer's Rule}}$$

- E.g. Gauss elimination scales like: $\mathcal{O}\left(\frac{n^3}{3} - \frac{n}{3}\right)$

Determinants

- Determinants can be evaluated efficiently using elimination methods.
- Consider an upper triangular matrix resulting from elimination:

$$\underline{\underline{U}} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots \\ 0 & u_{22} & u_{23} & \cdots \\ 0 & 0 & u_{33} & \\ 0 & 0 & 0 & \ddots \end{bmatrix}$$

$$\det(\underline{\underline{U}}) = (\text{expand along } 1^{\text{st}} \text{ column})$$

$$= u_{11}(-1)^{1+1}\left|\underline{\underline{M_{11}}}\right| = u_{11}(-1)^{1+1}\left[u_{22}(-1)^{1+1}\left|\underline{\underline{M_{22}}}\right|\right]$$

(expand along 1^{st} column of $\underline{\underline{M_{11}}}$)

$$= \ldots = \prod_i u_{ii}$$

- So evaluating dets of triangular matrices is trivial
- Performing simple elimination does not change the determinant, thus we can do fwd elimination then evaluate det from result.
- This is more efficient than the $(n-1)n!$ operations implied by cofactor expansions.
- Pivoting changes value of determinants in predictable ways:
  1) Scaling row by a const also scales det by const
  2) Swapping rows flips sign of det.

LU factorization

- One can factor a matrix in infinite ways; one useful way is to factor into upper & lower triangular matrices:

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

- This can be used to solve $\underline{\underline{A}}\,\underline{x} = \underline{b}$ as follows:

$$\underline{\underline{L}}\,\underline{\underline{U}}\,\underline{x} = \underline{b}\;; \qquad \underline{\underline{L}}^{-1}\underline{\underline{L}}\,\underline{\underline{U}}\,\underline{x} = \underline{\underline{L}}^{-1}\underline{b}$$

$$\underline{\underline{U}}\,\underline{x} = \underline{\underline{L}}^{-1}\underline{b}$$

- Defining $\underline{b}' = \underline{\underline{L}}^{-1}\underline{b}$ we have the following eqns:

(1) $\underline{\underline{L}}\,\underline{b}' = \underline{b}$ (2) $\underline{\underline{U}}\,\underline{x} = \underline{b}'$

- Thus we can view solving $\underline{\underline{A}}\,\underline{x} = \underline{b}$ as solving (1), (2) above
- This is advantageous since triangular systems can be solved easily using fwd/back substitution. Also for systems with multiple RHS, $\underline{\underline{L}}, \underline{\underline{U}}$ need only be computed once!
- An example with multiple RHS would be finding an inverse via LU factorization:

find $\underline{\underline{A}}^{-1}$ s.t. $\underline{\underline{A}}\,(\underline{\underline{A}}^{-1}) = \underline{\underline{I}}$ can be viewed as

$$\left.\begin{aligned} \underline{\underline{A}}\,\underline{x}_1 &= \underline{b}_1 \\ \underline{\underline{A}}\,\underline{x}_2 &= \underline{b}_2 \\ \underline{\underline{A}}\,\underline{x}_3 &= \underline{b}_3 \end{aligned}\right\} \quad \underline{\underline{A}}^{-1} = [\underline{x}_1\ \underline{x}_2\ \underline{x}_3]$$

$\underline{b}_1 = I_{i1}$ (first column of $\underline{\underline{I}}$)

$\underline{b}_2 = I_{i2}$ (2nd " " ")

$\underline{b}_3 = I_{i3}$ (3rd " " ")

Tridiagonal Systems

- Consider the system:

$$\underline{\underline{T}}\,\underline{x} = \underline{b};$$

$$\underline{\underline{T}} = \begin{bmatrix} T_{11} & T_{12} & & & & & \\ T_{21} & T_{22} & T_{23} & & & & \\ & T_{23} & T_{33} & T_{34} & & & \\ & & T_{34} & T_{44} & T_{45} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & T_{i,i-1} & T_{ii} & T_{i,i+1} & \\ & & & & \ddots & \ddots & \\ & & & & & T_{n,n-1} & T_{nn} \end{bmatrix}$$

- Elimination applied to this system is particularly simple. E.g. only one element in row 2 (column 1) to be eliminated using row 1.

$$R_2 \to R_2 - \frac{T_{21}}{T_{11}} R_1; \text{ viz. } T_{2j} = T_{2j} - \frac{T_{21}}{T_{11}} T_{1j}$$

$$j \in \{2,3\}$$

- Now fwd elimination can proceed downward to subsequent rows...

- We must also alter RHS from elim step:

$$b_2 = b_2 - \frac{T_{21}}{T_{11}} b_1$$

- These two steps can be encoded generally as (j^{th} row)

1) $$T_{ij} = T_{ij} - \left(\frac{T_{i,i-1}}{T_{i-1,i-1}} \right) T_{i-1,j}$$

$$\left(i \in \{1, n\} \; ; \quad j \in \{i-1, i+1\} \right)$$

there are two of these to perform:

$$T_{i,i} = T_{i,i} - \left(\frac{T_{i,i-1}}{T_{i-1,i-1}} \right) T_{i-1,i}$$

$$T_{i,i+1} = T_{i,i+1} - \left(\frac{T_{i,i-1}}{T_{i-1,i-1}} \right) T_{i-1,i+1}$$

2) Following by RHS:

$$b_i = b_i - \left(\frac{T_{i,i-1}}{T_{i-1,i-1}} \right) b_{i-1}$$

- Following these two steps $\underline{\underline{T}}$ has been triangularized

$\underline{\underline{T}} \rightarrow \underline{\underline{T}}'$; we can now perform backsubstitution...

$$x_n = b_n / T_{nn} \; ; \quad x_i = \frac{1}{T_{ii}} \left(b_i - T_{i,i+1} x_{i+1} \right)$$

Condition Number

- Elimination methods can be sensitive to round-off error which propagates thru subsequent steps...
- One way to measure this sensitivity is thru the matrix condition number which is computed using norms
- A norm is a measure of magnitude & satisfies the constraints:

a) $\|\underline{\underline{A}}\| > 0$

b) $\|\underline{\underline{A}}\| = 0$ iff $\underline{\underline{A}} = \underline{\underline{0}}$

c) $\|\alpha \underline{\underline{A}}\| = |\alpha| \, \|\underline{\underline{A}}\|$

d) $\|\underline{\underline{A}} + \underline{\underline{B}}\| = \|\underline{\underline{A}}\| + \|\underline{\underline{B}}\|$

e) $\|\underline{\underline{A}}\,\underline{\underline{B}}\| \leq \|\underline{\underline{A}}\| \, \|\underline{\underline{B}}\|$ (Schwarz inequality)

- Commonly used vector norms:

$$\|\underline{x}\|_1 = \sum_i |x_i| \qquad L^1\text{-norm}$$

$$\|\underline{x}\|_2 = \sqrt{\sum_i x_i^2} \qquad L^2\text{-norm} \quad \text{(Euclidean)}$$

$$\|\underline{x}\|_\infty = \max(|x_i|) \qquad L^\infty\text{-norm}$$

- Matrix norms can be similarly defined:

$$\|\underline{\underline{A}}\|_1 = \max_j \left\{ \sum_i |A_{ij}| \right\} \quad \text{(max col. sum)}$$

$$\|\underline{\underline{A}}\|_\infty = \max_i \left\{ \sum_j A_{ij} \right\} \quad \text{(max row sum)}$$

$$\|\underline{\underline{A}}\|_2 = \min(\lambda_i) \quad \text{(min eigenvalue, spectral norm)}$$

$$\|\underline{\underline{A}}\|_e = \sqrt{\sum_{i,j} A_{ij}^2} \quad \text{(Euclidean norm)}$$

- The condition number measures sensitivity to small changes (e.g. round-off) in system.
- Considering the system:

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

we know $\|\underline{b}\| \le \|\underline{\underline{A}}\|\,\|\underline{x}\|$

- Let us examine system perturbed by: $\underline{\delta b}$

$$\underline{\underline{A}}(\underline{x} + \underline{\delta x}) = \underline{b} + \underline{\delta b} = \underline{\underline{A}}\,\underline{x} + \underline{\underline{A}}\,\underline{\delta x} = \underline{b} + \underline{\delta b}$$

- Subtract original equation:

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

$$\Rightarrow \quad \underline{\delta x} = \underline{\underline{A}}^{-1}\,\underline{\delta b}$$

thus $\|\underline{\delta x}\| \le \|\underline{\underline{A}}^{-1}\|\,\|\underline{\delta b}\|$

- Multiplying the two Schwarz-inequalities:

$$\|\underline{b}\|\,\|\underline{\delta x}\| \le \|\underline{\underline{A}}\|\|\underline{x}\|\,\|\underline{\underline{A}}^{-1}\|\,\|\underline{\delta b}\|$$

$$\|\underline{b}\|\,\|\underline{\delta x}\| \le \|\underline{x}\|\,\|\underline{\delta b}\|\left(\|\underline{\underline{A}}\|\,\|\underline{\underline{A}}^{-1}\|\right)$$

$$\frac{\|\underline{\delta x}\|}{\|\underline{x}\|} \le \frac{\|\underline{\delta b}\|}{\|\underline{b}\|}\, C(\underline{\underline{A}})$$

$$C(\underline{\underline{A}}) \equiv \|\underline{\underline{A}}\|\,\|\underline{\underline{A}}^{-1}\| \quad \text{(condition number)}$$

- A large condition number $\Rightarrow$ modest $\|\underline{\delta b}\|/\|\underline{b}\|$ can lead to large $\|\underline{\delta x}\|/\|\underline{x}\|$ (BAD!!)
- $C(\underline{\underline{A}}) \sim 1$ is best.

Iterative Methods

- One drawback of elimination methods is the accumulation of error through successive steps.
- Iterative methods can address this issue, but require a matrix to be <u>diagonally dominant</u>, i.e.

$$\exists i \quad \text{s.t.} \quad |A_{ii}| \geq \sum_{j,\, j \neq i} |A_{ij}|$$

- Many systems will not meet this requirement...
- Iterative methods start with an initial guess and then refine that guess repeatedly until <u>convergence</u> is achieved.
- Consider: $\underline{\underline{A}}\,\underline{x} = \underline{b}$, for some guess $\underline{x}^{(0)}$

  the <u>residual</u> (error in a sense) is $\underline{b} - \underline{\underline{A}}\,\underline{x}^{(0)}$

$$R_i^{(k)} = b_i - \sum_j A_{ij} x_j^{(k)}$$

residual on i^{th} var @ iteration k

- When the residual is small the iterated solution is near the true solution.
- When the solution $\underline{x}^{(k)}$ does not change appreciably between iterations, it has <u>converged</u>.

- Jacobi iteration involves a correction to each variable based on residual:

$$x_i^{(k+1)} = x_i^{(k)} + \frac{R_i^{(k)}}{A_{ii}}$$

- The basis of this idea is that, given some guess for all vars. (sans i) we can directly compute the variable i:

$$A_{ii} x_i^{(k+1)} \approx b_i - \sum_{j \neq i} A_{ij} x_j^{(k)}$$

$$= A_{ii} x_i^{(k)} + b_i - \sum_j A_{ij} x_j^{(k)}$$

$$\therefore \quad x_i^{(k+1)} = x_i^{(k)} = \frac{1}{A_{ii}} \left(b_i - \sum_j A_{ij} x_j^{(k)} \right)$$

- Letting $\underline{\Delta x} = \underline{x}^{(k+1)} - \underline{x}^{(k)}$ we can define the convergence as:

$$|\underline{\Delta x}| \leq \varepsilon \quad \text{for some iteration step}$$

- Relative error can also be used for convergence but becomes problematic if a solution is near zero.

- Gauss-Seidel iteration uses the partially updated solution to determine $R_i^{(k)}$:

$$R_i^{(k)} = b_i - \sum_{j<i} A_{ij} x_j^{(k+1)} - \sum_{j \geq i} A_{ij} x_j^{(k)}$$

$$x_i^{(k+1)} = x_i^{(k)} + \frac{R_i^{(k)}}{A_{ii}} \quad \text{(same)}$$

- Convergence is almost always faster than Jacobi iteration & it's "simpler" to code.
- Successive over-relaxation (SOR) is a further refinement to speed convergence.

$$x_i^{(k+1)} = x_i^{(k)} + \omega \frac{R_i^{(k)}}{A_{ii}} \qquad (\omega < 2)$$

- The relaxation parameter (ω) must be chosen essentially by hand & some choices will make convergence slower...