Solution of Equations by Iteration

Method: to solve $f(x) = 0$, rewrite as $x = \theta(x)$. Make an initial guess x_0. Iterate: $x_{n+1} = \theta(x_n)$, until suitable convergence criteria are satisfied. Example: $e^x = 3x^2$. Several ways of rewriting this: $x = \log_e(3x^2)$; $x = \pm\sqrt{(1/3e^x)}$.

Cobweb diagrams: plot $y = x$ and $y = \theta(x)$ [$y = -\sqrt{(e^x/3)}$ in this example]. Start at $(x_0, 0)$. Draw a vertical line to $(x_0, \theta(x_0))$; draw a horizontal line to (x_1, x_1); draw a vertical line to $(x_1, \theta(x_1) = x_2)$, etc. Theorem: If $-1 < {}^{d\theta(x)}/_{dx} < 1$ near an intersection of $y = x$ and $y = \theta(x)$ (i.e. the roots of our equation), then the process converges. Generally a good iterative method is *Newton's method*, $X_{n+1} = X_n - [f(X_n)/f'(X_n)]$.

Numerical Linear Algebra

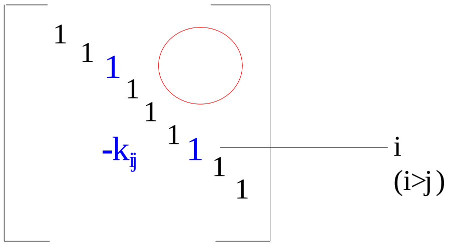
Problem: Solve $x + 3y + 4z = 3$, $2x + 5y + 7z = 5$, $5x + 8y + 9z = 16$ i.e. $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$. Method: Write $A = LU$, with L = *Lower Triangular* and U = *Upper Triangular*. Then, solve $L\underline{\mathbf{c}} = \underline{\mathbf{b}}$ by *forward* substitution and solve $U\underline{\mathbf{x}} = \underline{\mathbf{c}}$ by *backward* substitution.

$$A=\begin{bmatrix}1&3&4\\2&5&7\\5&8&9\end{bmatrix}\sim\begin{bmatrix}1&3&4\\0&-1&-1\\0&-7&-11\end{bmatrix}\begin{matrix}\\R_2:=R_2-2R_1\\R_3:=R_3-5R_1\end{matrix}\sim\begin{bmatrix}1&3&4\\0&-1&-1\\0&0&-4\end{bmatrix}\begin{matrix}\\ \\R_3:=R_3-7R_2\end{matrix}\quad .\ A=\begin{bmatrix}1&3&4\\2&5&7\\5&8&9\end{bmatrix}=\begin{bmatrix}1&0&0\\2&1&0\\5&7&1\end{bmatrix}\begin{bmatrix}1&3&4\\0&-1&-1\\0&0&-4\end{bmatrix}$$

Notes on the above: in the first array, 1 is the pivot (*position* 1,1). In the 2nd array, -1 is the pivot (*position* 2,2). The **numbers** 2, 5 and 7 in the descriptions *after* the arrays are the <u>subtraction coefficients</u>. And in the final **double** array, the *1st* array is the "L" array.

How Does This Happen? Every row subtraction can be performed by **left** multiplication by an "*elementary matrix*". $E_{ij}(m,n) = 1$; $E_{ij}(i,j) = -k_{ij}$; $E_{ij}(m,n) = 0$ otherwise. $E_{ij}A$ = matrix obtained **from** A using row operations $k_i := R_i - k_{ij} R_j$.

$E_{ij} =$



Example: $$E_{42}A=\begin{bmatrix}1&&&\\&2&&\\&&1&\\&-3&&1\end{bmatrix}\begin{bmatrix}a&b&c&d\\e&f&g&h\\p&q&r&s\\u&v&w&x\end{bmatrix}=\begin{bmatrix}a&b&c&d\\e&f&g&h\\p&q&r&s\\-3e+u&-3e+v&-3e+w&-3e+x\end{bmatrix}\ R_4:=R_4-3R_2$$

For a **typical** 3×3 matrix we perform $E_{32}\,E_{31}\,E_{21}\,A = U$:

$$\begin{bmatrix}1&&\\0&1&\\0&-7&1\end{bmatrix}\begin{bmatrix}1&&\\0&1&\\-5&0&1\end{bmatrix}\begin{bmatrix}1&&\\-2&1&\\0&0&1\end{bmatrix}=\begin{bmatrix}1&&\\0&1&\\0&-7&1\end{bmatrix}\begin{bmatrix}1&&\\-2&1&\\-5&0&1\end{bmatrix}=\begin{bmatrix}1&&\\-2&1&\\9&-7&1\end{bmatrix}.$$

The **picture** implies $E_{31}\,E_{21}\,A = E^{-1}_{32}\,U$. So $E_{21}\,A = E_{31}^{-1}\,E_{32}^{-1}\,U$. Now $A = E_{21}^{-1}\,E_{31}^{-1}\,E_{32}^{-1}\,U$. The R.H.S. is **equal** to L. So the required L is a product of inverse elementary matrices. The *inverse* of "subtract row J from row i k_{ij} times" is "add row j to row i k_{ij} times" (see over).

$$\text{i.e.}\begin{bmatrix} 1 & & & \\ & 1 & & \\ & -k_{ij} & 1 & \\ & & & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & +k_{ij} & 1 & \\ & & & 1 \end{bmatrix}$$

In *our* example, $L = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 5 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & +7 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ & & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ & 1 & \\ 5 & 7 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ 5 & 7 & 1 \end{bmatrix}.$

In **general**, L_{ij} = *coefficient* k_{ij} from $R_i := R_i - k_{ij}R_j$. Now *solve* L<u>**c**</u> = <u>**b**</u> (By *Fwd*. sub.).

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 7 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \\ 8 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 16 \end{bmatrix}$$

N.B. <u>**c**</u> is the *last column* which you would have got by performing **Gaussian elimination** in an enhanced matrix: $\begin{bmatrix} 3 \\ 5 \\ 16 \end{bmatrix} \; R_2 - 2R_1 \rightarrow \begin{bmatrix} 3 \\ -1 \\ 16 \end{bmatrix} \underset{R_3 - 5R_1}{\rightarrow} \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \underset{R_3 - 7R_2}{\rightarrow} \begin{bmatrix} 3 \\ -1 \\ 9 \end{bmatrix}$

Now solve U<u>**x**</u> = <u>**c**</u> (By *backward* substitution — working up from the *bottom* row).

$\begin{bmatrix} 1 & 3 & 4 \\ 0 & -1 & -1 \\ 0 & 0 & -4 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 8 \end{bmatrix}$ (Note: **2nd** array was <u>**x**</u>).

So the solution is x = 2, y = 3, z = -2. We have an **exact** solution having worked with *integers* throughout.

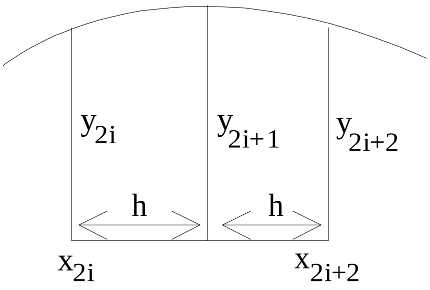
In each case the pivot was the *smallest* entry in the column. Computationally, to <u>minimise</u> round-off errors, make the pivot as **large** as possible.

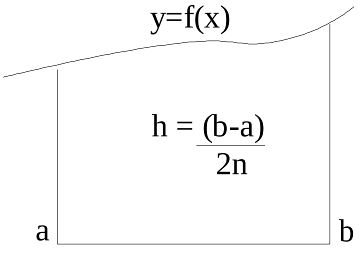
$$A\begin{bmatrix} 1 & 3 & 4 \\ 2 & 5 & 7 \\ 5 & 8 & 9 \end{bmatrix} \leftarrow R_1 \Leftrightarrow R_3 \rightarrow \begin{bmatrix} 5 & 8 & 9 \\ 2 & 5 & 7 \\ 1 & 3 & 4 \end{bmatrix} \begin{matrix} R_2 - \frac{2}{5}R_1 \\ R_3 - \frac{1}{5}R_1 \end{matrix} \text{. Then } |k_{ij}| \le 1 \sim \begin{bmatrix} 5 & 8 & 9 \\ 0 & ? & ? \\ 0 & ? & ? \end{bmatrix}.$$

Notes: With the **1st column** of A, find the element with the <u>greatest</u> modulus. In the last array, in the *2nd column*, find the largest of the two values represented as <u>question</u> marks, and then **continue** as before.

<u>Parabolic Approximation</u>

The **area** under the parabola (1st picture) is $(y_{2i}+4y_{2i+1}+y_{2i+2})(^h/_3)$. **Method** (2nd picture): Divide into 2n strips of width h. Fit a *parabola* to each pair of strips: $x_i = a+ih$, $y_i = f(x_i)$. $A = (y_0+4y_1+y_2)(^h/_3) + (y_2+4y_3+y_4)(^h/_3) + \ldots + (y_{2n-2}+4y_{2n-1}+y_{2n})(^h/_3) = (y_0+4y_1+2y_2+4y_3+2y_4+ \ldots +2y_{2n-2}+4y_{2n-1}+y_{2n})(^h/_3)$.





# strips	Trapezium	Simpson
8	1.4978	1.5418
16	1.5449	1.5606
32		1.5672

# strips	area	diff.	ratio
2^n	a_n		
2^{n+1}	a_{n+1}	d_{n+1} *1	
2^{n+2}	a_{n+2}	d_{n+2}	r_{n+} *2

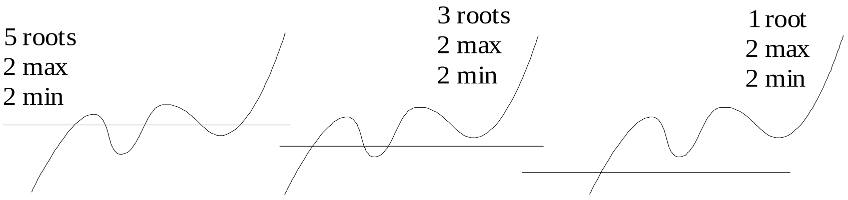
(($^\pi/_2$) = 1.5708). The *ratios of differences* tend to be approximately to the **power** of a half. Assume the ratio of differences are all approximately equal to r, say. This assumes that we repeatedly **double** the number of strips and obtain new *approximate* values.

Notes on table: *1: (a_{n+1})-(a_n). *2: $\frac{d_{n+2}}{d_{n+1}}$.

Assume the process **converges to give** area ‘a’ i.e. $a = \lim_{n\to\infty} a_n$. Also $a = a_n+d_{n+1}+d_{n+2}+\ldots$ We have *assumed that* $\frac{d_{n+2}}{d_{n+1}}, \frac{d_{n+3}}{d_{n+2}}$ = approximately r. i.e. $d_{n+2} = rd_{n+1}$; $d_{n+3} = rd_{n+2} = r^2d_{n+1}$; $d_{n+j} = r^{j-1}d_{n+1}$. So **a** *is approximately* $a_n + d_{n+1} + rd_{n+1} + r^2d_{n+1} + \ldots = a_n+d_{n+1}(1+r+r^2+\ldots)$ (*Geometrical progression*). The **extrapolation** formula is $\boxed{a = a_n + \frac{d_{n+1}}{(1-r)}}$. (Don’t know *r* so use r_{n+1}). *Alternative* formula: ‘a’ is approximately $a_{n+1} + \frac{r_{n+1}d_{n+1}}{(1-r_{n+1})}$.

Newton Iteration

Solution of e.g. $t^5+xt^3-yt^2-zt+1 = 0$. There are *several possible solutions*; some are shown in the diagram. Viewed as a **complex** equation, there are always *5 roots and 4 turning* points.



Newton **iteration** for x^4-1 = 0. Formula: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. For **our** function, $\phi(x) = x - \frac{(x^4-1)}{4x^3} = \frac{3x^4+1}{4x^3}$. The **four roots** of the equation are +1, -1, +i, and -i. If $x_0 = i\sqrt{2}$, then $\phi(x_0) = \frac{3(\sqrt{2})^4 i^4+1}{4(\sqrt{2})^3 i^3} = \frac{2i^4+1}{4\times 2\sqrt{2}(-i)} = \frac{12+1}{8\sqrt{2}\,i} = \frac{13}{\sqrt{2}\,i} = x_i$.