Non-Linear Methods

1. Introduction:

- Let $F: R - R$. We want to solve for x s.t. $F(x) = 0$. x is called a **root**.
- Examples of non-linear eqns:

1. $2x^2 + 7 = 0$
2. $x - e^{-x} = 0$ Transcendental Eqn

- Typically, we can't find a closed form or analytical soln to solve non-linear eqns. However, we can find iterative methods that generate an approximation.

I.e. Let $k = 0, 1, 2, \ldots$ As $k \to \infty$, $\hat{x}_k \to \tilde{x}$, where $\hat{x}_k$ is the approximation.

2. Fixed Point Methods (FPM):

- $F(\tilde{x}) = 0$ is called a **root finding problem.**
- $\tilde{x} = g(\tilde{x})$ is called a **fixed point problem.**
- $F(\tilde{x}) = 0$ is equivalent to $\tilde{x} = g(\tilde{x})$.

E.g. $\underbrace{x - e^{-x} = 0}_{\text{Root Finding problem}} \Leftrightarrow \underbrace{x = e^{-x}}_{\text{Fixed point problem}}$

- We can let $g(\tilde{x}) = \tilde{x} - F(\tilde{x})$ or let $g(\tilde{x}) = \tilde{x} - h(\tilde{x})F(\tilde{x})$ where $h(\tilde{x})$ is an auxiliary function.
- $g(\tilde{x}) = \tilde{x} - F(\tilde{x})$ is called the **first form.**
- $g(\tilde{x}) = \tilde{x} - h(\tilde{x})F(\tilde{x})$ is called the **second form.**

- If we use the first form, then $F(\tilde{x})=0$ is always equivalent to $\tilde{x}=g(\tilde{x})$.

Proof:
LHS: $F(\tilde{x})=0$
RHS: $\tilde{x}=g(\tilde{x})$

$$\begin{aligned} \tilde{x} &= g(\tilde{x}) \\ &= \tilde{x}-F(\tilde{x}) \\ 0 &= -F(\tilde{x}) \\ F(\tilde{x}) &= 0 \end{aligned}$$

Hence, LHS = RHS

- If we use the second form, if $F(\tilde{x})=0$, then $\tilde{x}=g(\tilde{x})$. However, we could have $\tilde{x}=g(\tilde{x})$ but $F(\tilde{x})\neq 0$. This situtation occurs if $h(\tilde{x})=0$. Hence, the two equations aren't equivalent. Furthermore, after we find a fixed point, we need to check if it is a root.
- The advantage of the second form is that there's flexibility in designing $g(\tilde{x})$, to make iteration converge faster.

3. Fixed Point Iteration (FPI):

- Start with an approximate soln $\hat{x}_0$ then iterate $\hat{x}_{k+1}=g(\hat{x}_k)$, $k=0,1,2,\ldots$ until convergence or failure.
- E.g. 1 Let $F(x)=x^2+2x-3$. We know that the roots are 1 and -3.

Consider the FPI $x_{k+1}=x_k+\dfrac{(x_k)^2+2x_k-3}{(x_k)^2-5}$

for the fixed point problem

$$\begin{aligned} x &= g(x) \\ &= x+\frac{x^2+2x-3}{x^2-5} \end{aligned}$$

We see that this is the second form of the fixed point problem where $h(x) = \frac{-1}{x^2-5}$.

Since $h(x) \neq 0, \forall x \in R$, this means that we don't need to check if a fixed point is a root.

If we start with $\hat{x}_0 = -5$, then $\hat{x}_k$'s approach -3.
If we start with $\hat{x}_0 = 5$, then $\hat{x}_k$'s do not converge.
If we start with $\hat{x}_0 = 0$, then $\hat{x}_k$'s converge to 1.

Hence, we can see that depending on $\hat{x}_0$, the FPI may converge to some fixed point or may not converge.

4. Fixed Point Theorem (FPT):

- If there's an interval $[a,b]$ s.t.

1. $g(x) \in [a,b] \ \forall x \in [a,b]$
2. $\|g'(x)\| \leq L < 1 \ \forall x \in [a,b]$

then $g(x)$ has a unique fixed point in $[a,b]$.

Proof:

Note: This proof has 3 parts but we were only shown part 1. The other parts were left for assignments.

Start with any initial guess, $\hat{x}_0 \in [a,b]$, and iterate.
$\hat{x}_{k+1} = g(\hat{x}_k)$, $k = 0, 1, 2, \ldots$
Then, all $\hat{x}_k \in [a,b]$.

Furthermore, $x_{k+1} - x_k = g(x_k) - g(x_{k-1})$
$= g'(n_k)(x_k - x_{k-1})$
for some $n_k \in [x_{k-1}, x_k]$
$\subset [a,b]$

We know this by the Mean Value Theorem (MVT)

Therefore, $|x_{k+1} - x_k| \leq |g'(n_k)(x_k - x_{k-1})|$
$= |g'(n_k)||x_k - x_{k-1}|$
$= L|x_k - x_{k-1}|$

Then, $|x_k - x_{k-1}| \leq \ldots \leq L^k |x_1 - x_0|$

Since we know that $L<1$, $|x_k - x_{k-1}| \to 0$ as $k \to \infty$.

This means that x_k converges to some point $\tilde{x} \in [a,b]$.

To complete the proof, we have to show 2 things:

1. $\tilde{x}$ is a fixed point. I.e. $\tilde{x} = g(\tilde{x})$.
2. $\tilde{x}$ is unique.

5. Rate of Convergence:

- Def: If $\lim\limits_{\tilde{x}_k \to \tilde{x}} \dfrac{|\tilde{x} - x_{k+1}|}{|\tilde{x} - x_k|^p} = c \neq 0$, then

we have the p-th order convergence to fixed point $\tilde{x}$.

E.g. 2 This example will show the importance of p. Consider the table of absolute errors of iterates, $|\tilde{x} - x_k|$, below.

k	$p=1,\ c=1/2$	$p=2,\ c=1$
0	10^{-1}	10^{-1}
1	$5 \cdot 10^{-2}$	10^{-2}
2	$2.5 \cdot 10^{-2}$	10^{-4}
3	$1.25 \cdot 10^{-2}$	10^{-8}
4	$6.125 \cdot 10^{-3}$	10^{-16}

We start with 10^{-1} for both systems.
For system 1, $p=1$, $c=1/2$, we get $|\tilde{x} - x_{k+1}| = \frac{|\tilde{x} - x_k|}{2}$. Hence, with each iteration, we divide by 2.

For system 2, $p=2$, $c=1$, we get $|\tilde{x} - x_{k+1}| = (|\tilde{x} - x_k|)^2$. Hence, with each iteration, we square.

Notice that despite having $c=1$, the third column converges much faster. This is because $p=2$ in the third column.

6. Rate of Convergence Thm:

– For the FPI $x_{k+1} = g(x_k)$, if $g'(\hat{x})$, $g''(\hat{x})$, $\ldots$, $g^{(p-1)}(\hat{x}) = 0$ but $g^{p}(\hat{x}) \neq 0$, then we have p-th order convergence.

Proof:

$$
\begin{aligned}
x_{k+1} &= g(x_k) \\
&= g(\tilde{x} + (x_k - \tilde{x})) \\
&= g(\tilde{x}) + g(x_k - \tilde{x})\, g'(\tilde{x}) + \frac{(x_k - \tilde{x})^2}{2!} g''(\tilde{x}) \\
&\quad + \ldots + \frac{(x_k - \tilde{x})^{p-1}}{(p-1)!} g^{(p-1)}(\tilde{x}) + \underbrace{\frac{(x_k - \tilde{x})^{p}}{p!} g^{p}(n_k)}_{\text{Remainder Term}}
\end{aligned}
$$

Taylor Series

If we have $g'(\tilde{x}), g''(\tilde{x}), \ldots, g^{(p-1)}(\tilde{x}) = 0$, then we get

$$x_{k+1} = g(\tilde{x}) + \frac{(x_k - \tilde{x})^{p}}{p!} g^{p}(n_k)$$

Recall that $g(\tilde{x}) = \tilde{x}$.

$$x_{k+1} = \tilde{x} + \frac{(x_k - \tilde{x})^{p}}{p!} g^{p}(n_k)$$

Rearranging the eqn, we get

$$\frac{x_{k+1} - \tilde{x}}{(x_k - \tilde{x})^{p}} = \frac{1}{p!} g^{p}(n_k)$$

As $k \to \infty$, $x_k \to \tilde{x}$, $n_k \in [\tilde{x}, x_k] \to \tilde{x}$.
We can rewrite this as

$$\lim_{x_k \to \tilde{x}} \frac{|x_{k+1} - \tilde{x}|}{|x_k - \tilde{x}|^p} = \frac{1}{p!} g^p(\tilde{x})$$

We see that if the p^{th} derivative of $g(\tilde{x})$ is not zero, we get p^{th} order convergence.

We can see that by using the second form of FPI, we can pick a $h(x)$ s.t. the p^{th} derivative of g is not zero.

7. Newton's Method:
- Formula: $x_{k+1} = x_k - \frac{F(x)}{F'(x)}$

This is the second form with $h(x) = \frac{1}{F'(x)}$.

- Suppose that $F(\tilde{x}) = 0$ and $F'(\tilde{x}) \neq 0$.

$$g'(x) = 1 - \left(\frac{F'(x)F'(x) - F(x)F''(x)}{(F'(x))^2} \right)$$

$$= \frac{(F'(x))^2 - (F'(x))^2 + F(x)F''(x)}{(F'(x))^2}$$

$$= \frac{F(x)F''(x)}{(F'(x))^2}$$

$$= 0$$

By the rate of convergence thm, Newton's Method has at least quadratic convergence for any function F.

- Geometric Interpretation of NM:

We want to solve $F(x)=0$ at an initial guess x_k which is an approximate/model to $F(x)$ by a linear polynomial $p(x)$ that satisfies the conditions:

$$\left.\begin{array}{l} 1.\ p(x_k) = F(x_k) \\ 2.\ p'(x_k) = F'(x_k) \end{array}\right\} \text{ p(x) is the tangent line to F(x).}$$

$$p_k(x) = F(x_k) + (x - x_k)F'(x_k)$$

Then, x_{k+1} is a root of $p_k(x)$.

I.e. $p_k(x_{k+1}) = 0 \Rightarrow F(x_k) + (x_{k+1} - x_k)F'(x_k) = 0$

$$\rightarrow x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)}$$

Note: NM doesn't always converge.

8. Secant Method:

- Take NM, $x_{k+1} = x_k - \dfrac{F(x_k)}{F'(x_k)}$

and approximate $F'(x_k)$ with $\dfrac{F(x_k) - F(x_{k-1})}{x_k - x_{k-1}}$

Now, the formula is $x_{k+1} = x_k - \dfrac{F(x_k)(x_k - x_{k-1})}{F(x_k) - F(x_{k-1})}$

- The secant method is not a FPI.
I.e. It can't be expressed as $x_{k+1} = g(x_k)$.
This is because there are 2 x_{k-1} terms.

- We can't directly use FCT or RCT to analyze this method. However, with some adjustments, we can prove that $p = \frac{1+\sqrt{5}}{2}$.

This is called superlinear convergence.
Note: NM has a $p = 2$.

- NM requires 2 function evaluations per iteration as F and F' aren't the same. The secant method only requires 1 function evaluation as we only need to compute $F(x_k)$. $F(x_{k-1})$ has already been computed in the previous step.
The effective rate of convergence takes into account cost per iteration as well as speed. Hence, the secant method is effectively faster than Newton's Method.

Note: The secant method doesn't always converge.

9. Bisection Method:

- Often, if we use NM/SM, we will start off with the wrong initial guess and won't get convergence.

- The Bisection Method will always guarantee convergence, but it's much slower.

- We need to find an $a<b$ s.t. $F(a) \leq 0 \leq F(b)$ or $F(b) \leq 0 \leq F(a)$. This means that there is at least 1 root in $[a,b]$.

Assume $F(a) \leq 0 \leq F(b)$
Loop until $b-a$ is small enough.
    Let $m = \frac{a+b}{2}$
    If $F(m) \leq 0$, let $a=m$, else $b=m$
    Repeat for the interval $[m,b]$ or $[a,m]$.

Then, we have linear rate of convergence with $p=1$, $c=1/2$, and with guaranteed convergence.

10. Hybrid Method:
- Combines slow, reliable methods with faster ones that require a more accurate guess.
E.g. Bisection + Newton

11. System of Non-Linear Eqn:
- Problem: Solve $F(\bar{x}) = \bar{0}$
- Newton's Method can be extended for a system of non-linear eqns.

$$\overline{x_{k+1}} = \overline{x_k} - \frac{F(\overline{x_k})}{F'(\overline{x_k})}$$ where F' is the Jacobian Matrix of F

When we divide by a matrix, we multiply by its inverse.

Hence, we get: $\overline{x_{k+1}} = \overline{x_k} - (F'(\bar{x}_k))^{-1} F(\bar{x}_k)$ or

$$\underbrace{(F'(\bar{x}_k))(\overline{x_{k+1}} - \overline{x_k}) = -F(\bar{x}_k)}_{\text{This is simply in the form of } A\bar{x} = \bar{b}.}$$

This is very expensive. We can use a pseudo NM by holding the Jacobian matrix fixed for a few iterations. This means that we can reuse the PA = LU factorization. This is alright so long as we are still converging.