Non-linear systems of equations / Rootfinding and Fixed Point Iteration

Monday, September 23, 2024 10:53 AM

Class 12: September 23, 2024

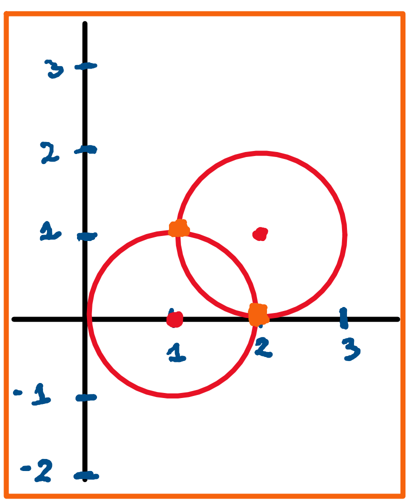
Recall: Last time, we wrapped up our first unit on rootfinding methods aimed at solving one non-linear equation in one real variable. At the very beginning of the unit, we argued that any non-linear equation like x + cos(x) = 3 can be easily turned into a problem of the form: *find x such that f(x) = 0.* We then spent some time developing and analyzing the following algorithms:

1. *Bisection method:* Assuming f is C[a,b] and changes sign, it is guaranteed to converge linearly with rate ~0.5 .
2. *Fixed Point Iteration:* If we can find g(x) such that g(x) = x is equivalent to f(x) = 0, we can apply FPI to find the roots of f(x). Assuming g is continuous, g(x) in [a,b] and is "contractive" around the root(s), it converges at least linearly (can be higher order, e.g. quadratic!). You can usually only get guarantees for x0 close enough to r.
3. *Newton-Raphson method:* Uses access to f'(x) to define a step as the root of the tangent line at the previous guess. It gives us a very special FPI method that, for x0 close enough to r, converges quadratically. Very powerful, but unreliable unless we can pair it with something else.
4. *Secant method:* Imitates the best features of Newton while avoiding computing f'(x). It uses two initial guesses x0 and x1 to define the secant line, the step is then the root of the secant line using the latest 2 iterations. For x0 close enough to r, converges superlinearly (order phi ~ 1.618).

However, it is rare to just have to solve one equation in one unknown. In applications, more often than not, we will have M non-linear equations with N unknowns, where M and N are potentially large. In this section, we will generalize some of the rootfinding methods discussed above to this case, and discuss how implementation / performance / costs behave in this setting. What is the same as the 1D case? What is different?

SYSTEMS OF NONLINEAR EQS:

EXAMPLE: Find the intersection pts for two circles of radius 1 centered at (1,0) and (2,1).



Two solutions:
$\vec{r_1} = (2,0)$, $\vec{r_2} = (1,1)$.

Algebraically:

$$\begin{cases} \text{Eq 1.} & (x_1-1)^2 + x_2^2 = 1 \\ \text{Eq 2.} & (x_1-2)^2 + (x_2-1)^2 = 1 \end{cases}$$

SOLUTIONS OF A SYSTEM OF 2 NONLIN EQS IN 2 VARS.

How do we relate this to rootfinding?
DEFINE $F: \mathbb{R}^2 \to \mathbb{R}^2$

$$F(\underbrace{[x_1, x_2]}_{\vec{x}}) = \begin{bmatrix} f_1([x_1,x_2]) \\ f_2([x_1,x_2]) \end{bmatrix} = \begin{bmatrix} (x_1-1)^2 + x_2^2 - 1 \\ (x_1-2)^2 + (x_2-1)^2 - 1 \end{bmatrix} = \begin{bmatrix} \vec{0} \\ \vec{0} \end{bmatrix}$$

s.t. solving our system $\Leftrightarrow$ finding $\vec{r}$ such that $\underline{F(\vec{r}) = \vec{0}}$.

In general: $F: \mathbb{R}^n \to \mathbb{R}^m$, $F(\underbrace{\vec{x}}_{\in \mathbb{R}^n}) = \underbrace{\vec{0}}_{\mathbb{R}^m}$

m equations $\underline{f_i(\vec{x}) = 0}$, $i = 1, \dots, m$.
in n variables

Assume $m = n$.

FIXED POINT ITERATION

Find G such that $G(\vec{x}) = \vec{x} \longleftrightarrow F(\vec{x}) = \vec{0}$.

Candidates:

$$G(\vec{x}) = \vec{x} + c\,F(\vec{x}), \quad c \neq 0.$$

$$G(\vec{r}) = \vec{r}$$

$$\cancel{\vec{r}} = \cancel{\vec{r}} + c\,F(\vec{r}) \to cF(\vec{r}) = \vec{0}.$$

$$G(\vec{x}) = \vec{x} + C(\vec{x})F(\vec{x})$$

$$\begin{bmatrix} g_1(x_1,x_2) \\ g_2(x_1,x_2) \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + C(\vec{x}) \begin{bmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{bmatrix}$$

$$G(\vec{x}) = \vec{x} + \underbrace{\mathbb{C}(\vec{x})}_{2\times 2 \text{ matrix.}} F(\vec{x})$$

$$\vec{r} = G(\vec{r}) = \vec{r} + \mathbb{C}(\vec{r})F(\vec{r})$$

$$\boxed{\vec{0} = \underbrace{\mathbb{C}(\vec{r})}_{2\times 2}\underbrace{F(\vec{r})}_{2\times 1}}$$

$\mathbb{C}(\vec{x})$ invertible around $\vec{r}$

$\left| \quad \vec{r} \quad \right|$

1D: Condition $|g'(x)| \le K < 1$ on a neigh of r $\to$ Linear convergence.

nD: Condition on $\|J_G\| \le k < 1$?

- for $x, y \in [a,b]$, $|g(x) - g(y)| \le K|x-y|$, $K<1$

for $\vec{x}, \vec{y}$ on neigh of r, $\|G(\vec{x}) - G(\vec{y})\| \le K\|\vec{x} - \vec{y}\|$

$$\|(x_1, x_2)\|_2 = \sqrt{x_1^2 + x_2^2}$$

$$\|(x_1, x_2)\|_\infty = \max(|x_1|, |x_2|)$$

$$\|(x_1, x_2)\|_1 = |x_1| + |x_2|$$

$\vec{x}_n$ close to r,

1D: $e_{n+1} = g'(r)\, e_n + \frac{1}{2} g''(\xi_n) \cdot e_n^2$

nD: $\underline{\vec{e}_{n+1} = J_G(\vec{r}) \cdot \vec{e}_n} + \text{ H.O.T } O(\|\vec{e}_n\|^2)$