Nonlinear systems: Fixed point

Friday, September 22, 2023 2:06 PM

Warm-up

Given a function $f(x)$ & the goal of finding a fixed pt. What requirement do we have in gaurantee convergence to the unique fixed pt?

Soln: We need $|f'(x)| < 1$ in a neighborhood of the fixed pt & our intial guess should be in that neighborhood.

Non-linear systems of equations: Fixed pt

$f(x,y)$ & $g(x,y)$ what is a fixed pt of the non-linear set of functions

$$\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} ?$$

Soln: (α, β) is a fixed pt if

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} f(\alpha, \beta) \\ g(\alpha, \beta) \end{bmatrix}$$

The fixed pt iteration $\qquad x_{n+1} = f(x_n)$

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{bmatrix} \quad \text{for} \quad n = 0, 1, 2, \ldots$$

How do we know when this will converge?

Answer: follow 1D initiation.
Assume f & g are analytic (Nice enough!)

let $\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} \alpha + \Delta x_n \\ \beta + \Delta y_n \end{bmatrix}$ where Δx_n & Δy_n are errors at iteration n.

So our fixed pt iteration

$$\begin{bmatrix} \alpha + \Delta x_{n+1} \\ \beta + \Delta y_{n+1} \end{bmatrix} = \begin{bmatrix} f(\alpha + \Delta x_n, \beta + \Delta y_n) \\ g(\alpha + \Delta x_n, \beta + \Delta y_n) \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{\text{fixed pt}} + \underbrace{\begin{bmatrix} \Delta x_{n+1} \\ \Delta y_{n+1} \end{bmatrix}}_{\text{error}} = \underbrace{\begin{bmatrix} f(\alpha, \beta) \\ g(\alpha, \beta) \end{bmatrix}}_{= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}} + \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} \begin{bmatrix} \Delta x_n \\ \Delta y_n \end{bmatrix}$$

+ higher order terms

By definition of the fixed pt.

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \underbrace{\begin{bmatrix}f_x & f_y\\ g_x & g_y\end{bmatrix}\Bigg|_{(\alpha,\beta)}}_{G = \text{The Jacobian}} \begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix} + \text{higher order terms.}$$

$$\Rightarrow \begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} = \begin{bmatrix}f_x & f_y\\ g_x & g_y\end{bmatrix}\Bigg|_{(\alpha,\beta)} \begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix} + \underbrace{\text{H.O.T.}}$$

Drop the higher order terms because they are small compared to the 1st order term

our iteration is roughly

$$\begin{bmatrix}\Delta x_{n+1}\\ \Delta y_{n+1}\end{bmatrix} \doteq G\Big|_{(\alpha,\beta)} \begin{bmatrix}\Delta x_n\\ \Delta y_n\end{bmatrix}$$

$$= G^n\Big|_{(\alpha,\beta)} \begin{bmatrix}\Delta x_0\\ \Delta y_0\end{bmatrix}$$

Cliffnotes

$$\Delta x_{n+1} = f'(\alpha)\Delta x_n$$

$$= (f'(\alpha))^n \Delta x_0$$

Sidenote: $G^2 x = G(Gx)$

Under what conditions is this going to converge?
(looking for properties of the Jacobian)

let $G \in \mathbb{R}^{m \times m}$ (Here $m=2$) Assume that G is full rank.

- eigenvectors are linear independent.
  - form a basis for $\mathbb{R}^m$
- there are no 0 eigenvalues

let $\{(\lambda_j, v_j)\}_{j=1}^m$ denote eigenpairs of G.

This means that $G v_j = \lambda_j v_j$

Since $\{v_j\}_{j=1}^n$ are basis for $\mathbb{R}^m$, we can express any $w \in \mathbb{R}^m$ as

$$w = \sum_{j=1}^m a_j v_j \quad \text{for some constants } \{a_j\}_{j=1}^m$$

So what happens when we apply G to w

$$G w = G\left(\sum_{j=1}^m a_j v_j\right) = \sum_{j=1}^m a_j G v_j = \sum_{j=1}^m a_j \lambda_j v_j$$

$$\Rightarrow G^n w = \sum_{j=1}^m a_j \lambda_j^n v_j$$

For what values of λ_j does this converge to 0?

For what values of λ_j does this converge to 0?

Answer: $|\lambda_j| < 1 \quad \forall j$

Really we want $\begin{bmatrix} \Delta x_{n+1} \\ \Delta y_{n+1} \end{bmatrix} \approx G^n \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \to \bar{0}$

This will happen when $|\lambda_j| < 1 \quad \forall j$

In practice we look at convergence in norms not vector.

$$\left\| \begin{bmatrix} \Delta x_{n+1} \\ \Delta y_{n+1} \end{bmatrix} \right\| = \left\| G^n \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \right\| \leq \|G\|^n \left\| \begin{bmatrix} \Delta x_0 \\ \Delta y_0 \end{bmatrix} \right\|$$

It is sufficient for $\|G\| < 1$.

In practice we tend to use the 2-norm
ℓ_2-norm.

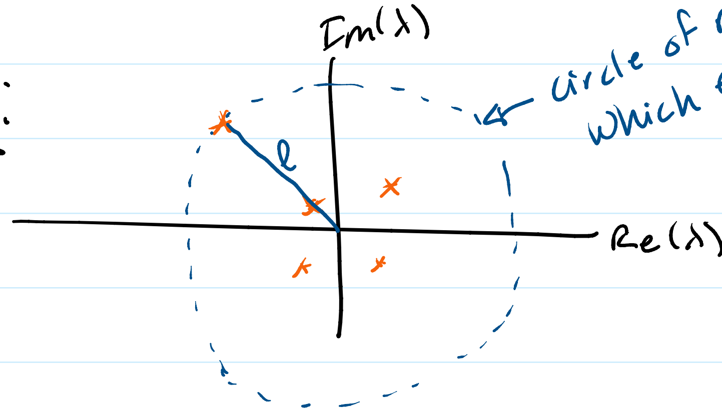
Thm 10.6 in the text has a more formal Statement.

Thm: $\|G\| \geq \rho(G)$

$\rho(G)$ is defined as the spectral radius of the matrix.

$$\rho(G) = \max_j |\lambda_j|$$

Cartoon:



circle of radius ρ which encloses all λ_j

$* = \lambda_j$

Ex:

(a) Can we use the fixedpt iteration for this system of functions?

$$f(x,y) = 3x$$
$$g(x,y) = \tfrac{1}{2}y$$

Make the Jacobian $G = \begin{bmatrix} 3 & 0 \\ 0 & \tfrac{1}{2} \end{bmatrix}$

$$\underbrace{\rho(G) = \max |\lambda_j|}_{\text{Check for } G} = 3 > 1$$

want the $\|G\|$ or $\rho(G) < 1$

No. the fixed pt iteration will not converge for any intial guess that is not $(x_0, y_0) = (0,0)$

Warm-up: What is the goal of fixed pt method for Non-linear systems?

Given a vector function $\bar{F}(\bar{x}) \in \mathbb{R}^m$ $\quad x \in \mathbb{R}^m$

Goal: Find $\bar{\alpha} \in \mathbb{R}^m$ st

$$\bar{F}(\bar{\alpha}) = \bar{\alpha}$$

(last time $m=2$ $\quad \bar{F}(x,y) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$

Fixed pt iteration

$$\bar{x}_{n+1} = \bar{F}(\bar{x}_n)$$

When does this iteration converge?

When $\rho(J) = \max_j |\lambda_j| < 1$

where J is the Jacobian evaluated at the fixed pt.

Pseudocode: Fixed pt for non-linear systems

Input: $\bar{F}(\bar{x})$ - vector function

$\bar{x}_0$ - initial guess

Input: $F(\bar{x})$ - vector function
$\bar{x}_0$ - initial Guess
tol - tolerance
N_{max} - max # of iterations.
(Optional: Jacobian: $J(\bar{x})$)

output:
$\bar{x}^*$ = approximate fixed pt
ier = error message

Steps:

Step1: $F_{x_0} = \bar{F}(\bar{x}_0)$

Step2: if $\|\bar{x}_0 - F_{x_0}\| == 0$

$x^* = x_0$
ier = 0
return.

Step3: For $j = 1, \cdots, N_{max}$

$\bar{x}_1 = \bar{F}(\bar{x}_0)$

if $\|\bar{x}_1 - \bar{x}_0\| < tol$ $\left(\text{or } \frac{\|\bar{x}_1 - \bar{x}_0\|}{\|x_1\|} < tol \right)$
$x^* = x_1$
ier = 0

return

$x_0 = x_1$

Step 4: $x^* = x_1$

ier=1

return

Difference from 1D:
- Norms instead of absolute value
- evaluate vector function F
- Norm is normally l_2