Systems of Nonlinear Eqs.

A. Donev, Spring 2021

Consider now a system of two nonlinear eqs. in 2 variables:

$$\begin{cases} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{cases}$$

Imagine we are given a guess for the solution $x_1^{(k)}$ and $x_2^{(k)}$ and we want to improve it to get $x_{1,2}^{(k+1)}$ (Worksheet 2)

There is no equivalent of bisection in higher dimensions!

1D is very special because the real line is ordered!
But the plane (or $\mathbb{R}^{d>1}$) is not ordered.

Newton's method, however, generalizes easily since it is based on a Taylor series!

$$f_1(x_1, x_2) = f_1(x_1^{(k)}, x_2^{(k)}) +$$

$$\frac{\partial f_1}{\partial x_1}(x_1^{(k+1)} - x_1) +$$

$$\frac{\partial f_1}{\partial x_2}(x_2^{(k+1)} - x_2)$$

and similarly for f_2

Write this using matrix notation as

$$\vec{f} = (f_1, f_2)^T$$

$$\vec{x} = (x_1, x_2)^T$$

$$\vec{f}(\vec{x}) = \vec{0}$$

$$f(\vec{x}^{(k+1)}) = f(\vec{x}^{(k)}) + \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 \end{bmatrix} \begin{bmatrix} x_1^{k+1} - x_1^{k} \\ x_2^{k+1} - x_2^{k} \end{bmatrix}$$

$$\vec{f}(\vec{x}^{(k+1)}) = \vec{f}(\vec{x}^{(k)}) + \overleftrightarrow{J}(\vec{x}^{(k)})(\vec{x}^{(k+1)} - \vec{x}^{(k)})$$

$$J_{ij} = \frac{\partial f_i}{\partial x_j}$$

Jacobian matrix

Idea in Newton's method:
Replace $\vec{f}(\vec{x})$ by its first-order Taylor series and find root of that linear function:

$$\vec{f}(\vec{x}^{(k+1)}) = \vec{0}$$

$$\left\{ \begin{aligned} &\overleftrightarrow{J}(\vec{x}^{(k)})\, \Delta x^{(k)} = -f(\vec{x}^{(k)}) \\ &\text{Newton's method} \\ &x^{(k+1)} = x^{(k)} + \Delta x^{(k)} \end{aligned} \right.$$

Linear system to be solved each iteration using LU factorization. Books may write

$$x^{(k+1)} = x^{(k)} - \cancel{\left(J^{(k)}\right)^{-1}} f^{(k)}$$ but

use backslash instead

Note: Linear systems are at the heart/core of NA!

For scalar functions of many variables

$$f : \mathbb{R}^n \to \mathbb{R}$$

Taylor series to second order is

$$f(\vec{x}+\Delta\vec{x}) = f(\vec{x}) + \vec{g}^T \Delta\vec{x} + \frac{1}{2}(\Delta\vec{x})^T \overleftrightarrow{H}\,\Delta\vec{x} + O(\|\Delta x\|^3)$$

where gradient (transpose of Jacobian) is column vector

$$\vec{g} = \nabla_{\vec{x}}\vec{f} = \vec{\nabla} f = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right]^T$$

and Hessian matrix is

$$\overleftrightarrow{H} = \overleftrightarrow{\nabla}_x^2 f$$

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} = H_{ji}$$

if f is twice continuously differentiable.

This is one of the reasons symmetric matrices are important!

For $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$

gradient (vector) replaced by Jacobian (matrix)

and

Hessian (matrix) replaced by rank-3 tensor

Convergence of Newton method

Let $\vec{f}(\vec{x}) = \vec{0}$

$\vec{x}$ ← root

$\vec{e}^{(k)} = \vec{x}^{(k)} - \vec{x} =$ error at step k

Near the root

$$\overleftrightarrow{J}(\vec{x}^{k}) \simeq \overleftrightarrow{J}(\vec{x})$$

so evaluate derivatives at <u>root</u> in analysis

$$\vec{f}(\vec{x}^{(k)}) = \vec{f}(\vec{x}) + \overleftrightarrow{J}(\vec{x})\vec{e}^{k} + \frac{1}{2}(\vec{e}^{k})^T \overleftrightarrow{H} (\vec{e}^{(k)})$$

$$\vec{e}^{(k+1)} = \vec{x}^{(k+1)} - \vec{x} = \overleftrightarrow{J}^{-1}\vec{f}(\vec{x}^{(k)})$$

$$= (\vec{x}^{(k)} - \vec{x}) - \overleftrightarrow{J}^{-1}\vec{f}(\vec{x}^{(k)})$$

$$\Rightarrow \quad \vec{e}^{\,(k+1)} = \vec{e}^{\,(k)} - \overleftrightarrow{J}^{-1} \vec{f}(\vec{x}^{\,(k)})$$

$$= \vec{e}^{\,(k)} - \overleftrightarrow{J}^{-1} \Big[\cancelto{0}{\vec{f}(\vec{x})} + \overleftrightarrow{J}\, \vec{e}^{\,(k)} + \frac{1}{2} (\vec{e}^{\,(k)})^T \overleftrightarrow{H} (\vec{e}^{\,(k)}) \Big]$$

$$\Rightarrow \vec{e}^{\,(k+1)} = -\frac{1}{2} (\overleftrightarrow{J}^{-1}) (\vec{e}^{\,(k)})^T \overleftrightarrow{H} (\vec{e}^{\,(k)})$$

rank 3 tensor

$$\|\vec{e}^{\,(k+1)}\| \leq \frac{\|\overleftrightarrow{J}^{-1}\| \, \|\overleftrightarrow{H}\|}{2} \|e^k\|^2$$

quadratic convergence

as long as Jacobian is not (nearly) singular

Compare this to 1D analysis

$$e^{(k+1)} \approx -\frac{1}{2}\frac{f''(x)}{f'(x)}\left(e^{(k)}\right)^2$$

So <u>if</u> Newton's method in higher dims converges, it will eventually converge fast once it gets sufficiently close to the root.

But Newton's method requires a good guess to converge (and since no bisection there is no easy way to get a good initial guess).