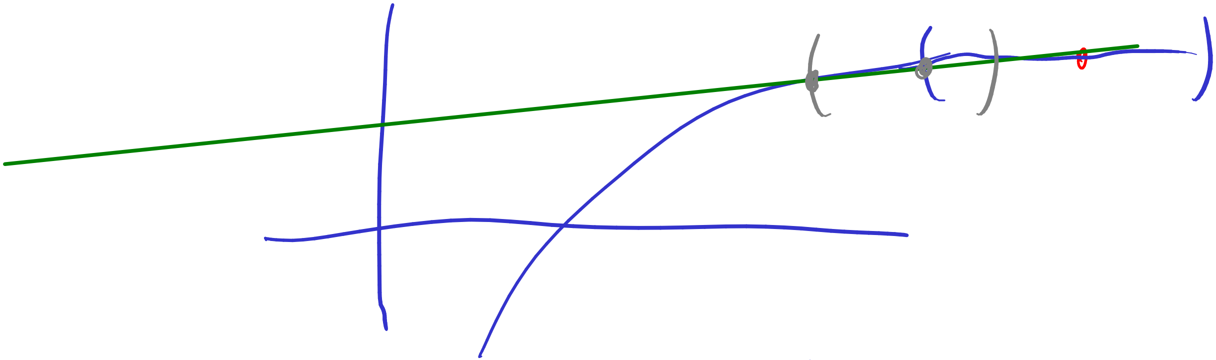
‘Trusting’ Newton and Secant

- The linear approximations in Newton and Secant are only good locally. How could we use that?



"Trust region"

"Hybrid": bracketing + superlinear

5.3 Methods in n Dimensions ("Systems of Equations")

$$f(\vec{x}) = f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x^2 y \sqrt{x}}{y} \\ x^2 y \sin(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- FPI
- Newton
- Secant

Fixed Point Iteration

$$\begin{aligned} \boldsymbol{x}_0 &= \langle \text{starting guess} \rangle \in \mathbb{R}^n \\ \boldsymbol{x}_{k+1} &= \underline{\boldsymbol{g}}(\boldsymbol{x}_k) \end{aligned}$$

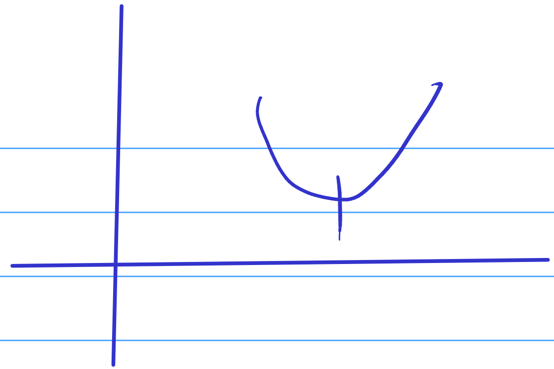
$x^* = g(x^*)$

$g(x) = f(x) + x$

- When does this converge?

$$J_g : \mathbb{R}^{n \times n} = \begin{pmatrix} \partial_1 g_1 & \cdots & \partial_n g_1 \\ \vdots & & \vdots \\ \partial_1 g_n & \cdots & \partial_n g_n \end{pmatrix} \qquad g = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}$$

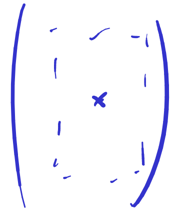
nd	1d
Will converge linearly if $\rho(J_g) < 1$.	$\lvert g'(x^*) \rvert < 1$
— " — quadratically if $J_g(x^*) = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$	$g'(x^*) = 0$



Newton's method

- What does Newton's method look like in *n* dimensions?

$$\vec{f}(\vec{x}+\vec{s}) \approx \vec{f}(\vec{x}) + J_f(\vec{x})\vec{s}$$



$$\vec{f}(\vec{x}) + J_f(\vec{x})\vec{s} \overset{!}{=} 0$$

$$J_f(\vec{x})\vec{s} = -\vec{f}(\vec{x})$$

Newton:

$\vec{x}_0 =$ (some starting guess)

$$\vec{x}_{k+1} = \vec{x}_k + \vec{s} = \vec{x}_k - J_f(\vec{x})^{-1} f(\vec{x})$$

$$x_{k+1} = x_k - \frac{f(x)}{f'(x)}$$

Demo: Newton's method in n dimensions

Secant in *n* dimensions?

- What would the secant method look like in *n* dimensions?

$$\underbrace{\vec{f}(\vec{x}+\vec{s})}\approx\underbrace{\vec{f}(\vec{x})}+\underbrace{J_f(\vec{x})}_{\substack{\text{unknown}\\ n^2}}\vec{s}$$

n equations

Broyden's method

6 Optimization

Optimization

- State the problem.

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

Want: $\vec{x}^*$

$$f(\vec{x}^*) = \min_{\vec{x}} f(\vec{x}) \quad \text{subject to} \quad \overbrace{\begin{array}{l} \vec{g}(\vec{x}) = 0 \\ \vec{h}(\vec{x}) \leq 0 \end{array}}^{\text{constraints}}$$

f, g, h are linear: "linear programming"

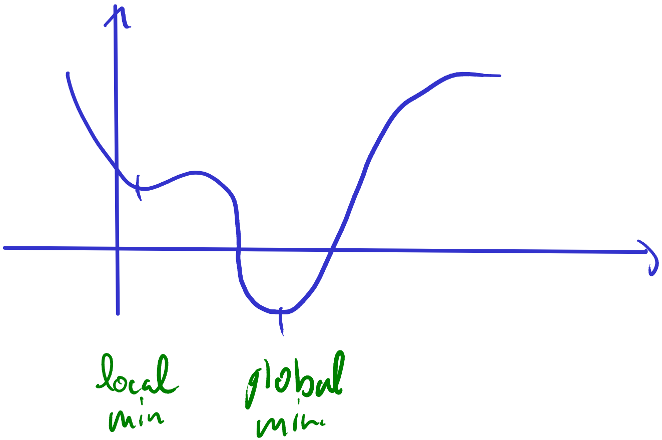
nonlinear " nonlinear — " —

$\|Ax - b\|$ $\to$ Another source of opt. problems:

$$\vec{\varphi}(\vec{x}) \approx 0 \qquad \min_{\vec{x}} \underbrace{\|\varphi(\vec{x})\|_2}_{f(\vec{x})}$$

Existence/Uniqueness

- Under what conditions on f can we say something about existence/uniqueness?



$$A = U \Sigma V^T$$

$$A^T A = U^T \Sigma^2 U$$

Optimality Conditions

- If we have found a candidate $\boldsymbol{x}^*$ for a minimum, how do we know it actually is one?
  To make this doable, assume f is smooth–i.e. has as many derivatives as needed.

Sensitivity and Conditioning

- How does optimization react to a slight perturbation of the minimum?

6.1 Methods for unconstrained opt. in one dimension