April 2, 2021

★ $\nabla f(x) = g(x) = 0$

Newton (1-D): $0 = g(x+h) \approx g(x) + g'(x)h \Rightarrow h = -g(x)/g'(x)$

$x_{k+1} = x_k - g(x_k)/g'(x_k)$

Taylor for multivariate: $g: \mathbb{R}^n \to \mathbb{R}^m$

$x \in \mathbb{R}^n$: $g(x) = 0$

$g(x+p) = g(x) + J(x)p + O(\|p\|^2)$ $\quad J(x) \in \mathbb{R}^{m \times n}$ $\quad J_{ik} = \frac{\partial [g(x)]_i}{\partial x_k}$

$0 = g(x+p) \approx g(x) + J(x)p \iff p = -J^{-1}(x)g(x)$

$m = n$ (J nonsing)

$x_{k+1} = x_k - J(x)^{-1} g(x)$

$g = \nabla f \Rightarrow J_{ik} = \frac{\partial g_i}{\partial x_k} = \frac{\partial f}{\partial x_i \partial x_k} = H$ $\quad J(x) = H(x)$

Newton for opt: $x_{k+1} = x_k - H(x)^{-1} \nabla f(x)$ $\quad \min_{x \in \mathbb{R}^n} f(x)$

Last time: $f(x) = \frac{1}{2} x^T A x + x^T b + c \qquad (A = A^T)$

$$\nabla f(x) = Ax + b \qquad H(x) = A$$

Newton:
$$\begin{aligned} x_{k+1} &= x_k - H(x_k)^{-1} \nabla f(x_k) \\ &= x_k - A^{-1}(Ax_k + b) \\ &= x_k - x_k - A^{-1} b \\ &= -A^{-1} b \end{aligned}$$

$$f(x) = \frac{1}{2} x^T A x - x^T b$$

$$\nabla f(x) = Ax - b = 0 \iff Ax = b$$

$$x_1 = A^{-1} b$$

Local convergence

Newton: $x_{k+1} \overbrace{- x_*}^{} = x_k - x_* - H_k^{-1} \nabla f_k - H_k^{-1} \nabla f_*$

$$e_{k+1} = e_k - H_k^{-1} (\nabla f_k - \nabla f_*) = H_k^{-1} (H_k e_k - (\nabla f_k - \nabla f_*))$$

$\int_0^1$

$\hookrightarrow$ By Lip. cont. $\leq L \| x_k - t e_k - x_k \|$