Lecture 7 — Newton's Method

Thursday, 20 January 2022 2:04 PM

Recall: Ex.1

$$\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}) := \|A\underline{x} - \underline{b}\|^2$$

No constraints

$A_{m\times n}$, $\underline{x}_{n\times 1}$, $\underline{b}_{m\times 1}$
rank $A = n$
$m \geq n$

FONC $\nabla f(\underline{x}^*) = 0$

Normal equations $(A^TA)\,\underline{x}^* = A^T\underline{b}$.

Given measurement points, we tried fit a 'straight line' that best fits the given data.

$\boxed{\text{Ex.2}}$ Given 'm' measurement of a process at 'm' points in time, fit a 'sinusoidal' curve to fit the measurement data.

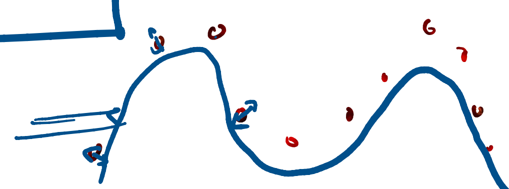
$$y = A \sin(\omega t + \phi)$$

A, ω, ϕ are unknown parameters.

$$\underline{x} = \begin{bmatrix} A \\ \omega \\ \phi \end{bmatrix}$$

Given

i	1	2	.....	m
t_i	t_1	t_2	....	t_m
y_i	y_1	y_2	.....	y_m



Find $\underline{x} = \begin{bmatrix} A \\ \omega \\ \phi \end{bmatrix}$ such that the objective function

$$f(\underline{x}) = \|\underline{r}(\underline{x})\|^2 = \underline{r}(\underline{x})^T \underline{r}(\underline{x})$$

$r_i(\underline{x}) = \underline{y_i - A\sin(\omega t_i + \phi)}$

$$\underline{r}(\underline{x}) = \begin{bmatrix} r_1(\underline{x}) \\ r_2(\underline{x}) \\ \vdots \\ r_m(\underline{x}) \end{bmatrix}$$

is a minimum.

Problem $\min \underline{r}(\underline{x})^T \underline{r}(\underline{x})$ (Non-linear least-square

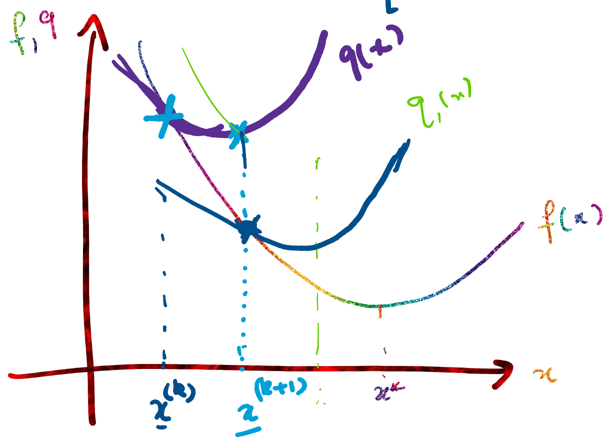
Problem $\min_{\underline{x} \in \mathbb{R}^n} \underline{g}(\underline{x})^T \underline{g}(\underline{x})$ (Non-linear least-square problem)

Unconstrained optimization Problem.

Newton's Method [Newton, Newton-Raphson, Newton - Simpson]

$\boxed{\min_{\underline{x} \in \mathbb{R}^n} f(\underline{x})}$ $\rightarrow \in C^{(2)}$

Root finding $g(x) = 0$.



Idea: 1) Given a starting point, construct a quadratic approximation to the objective function that matches with the value of the function, the first derivative and the second derivative at the point.

2] Minimize the approximate quadratic function instead of the objective function.

3] Use the minimizer of the approximate quadratic function as the starting point of the next step and repeat the process (iteratively).

$\boxed{1D}$ $f : \mathbb{R} \to \mathbb{R}$.
$q(x)$

$$\left.\begin{aligned} q(x^{(k)}) &= f(x^{(k)}) \\ q'(x^{(k)}) &= f'(x^{(k)}) \\ q''(x^{(k)}) &= f''(x^{(k)}). \end{aligned}\right]$$

$$q(x) = q(x^{(k)}) + (x - x^{(k)})\, q'(x^{(k)}) + \frac{(x - x^{(k)})^2}{2!}\, q''(x^{(k)})$$

$$q(x) = f(x^{(k)}) + (x - x^{(k)})\, f'(x^{(k)}) + \frac{(x - x^{(k)})^2}{2!}\, f''(x^{(k)}).$$

x^* is a minimizer for $q(x)$ $\Rightarrow$ $\boxed{\nabla q(x^*) = 0.}$

x^* is a minimizer for $q(x)$ $\Rightarrow$ $\boxed{\nabla q(x^*) = 0}$

$$\nabla q(x) = f'(x^{(k)}) + (x - x^{(k)}) f''(x^{(k)}) = 0$$

$$\Rightarrow (x^* - x^{(k)}) f''(x^{(k)}) = -f'(x^{(k)})$$

$$\Rightarrow x^* = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})} \quad (\neq 0)$$

$$\boxed{x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}}$$

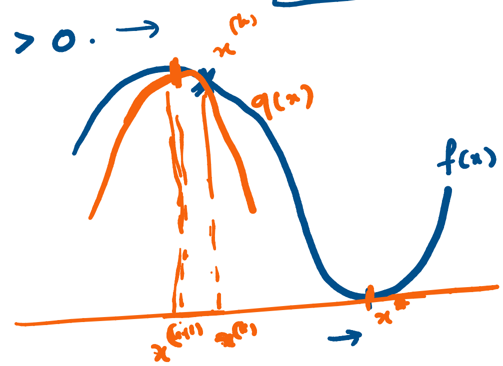
TOL

$\|x^{(k)} - x^*\| \le 10^{-5}$

$\|x^{(k)} - x^{(k+1)}\| \le 10^{-5}$

__Caveats__ Works well mostly if $f''(x) > 0$.

Fails with bad starting points



"__Quadratic convergence__"

__Root-finding__ $\boxed{g(x) = f'(x)}$

$$g(x^{(k+1)}) \approx g(x^{(k)}) + g'(x^{(k)})(x - x^{(k)})$$

$$\boxed{x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}}$$

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

$f(\underline{x}) \approx q(\underline{x})$ q is a quadratic approximation.

$q(\underline{x}^{(k)}) = f(\underline{x}^k)$

$\nabla q(\underline{x}^{(k)}) = \nabla f(\underline{x}^{(k)})$

$$\left.\begin{aligned}\nabla q(x^{(k)}) &= \nabla f(x^{(k)})\\ \nabla^2 q(x^{(k)}) &= \nabla^2 f(x^{(k)}) = F(x^{(k)})\end{aligned}\right]$$

FONC $\qquad \nabla q(x^*) = 0$

$$f(x) \approx f(x^{k}) + \underline{(x - x^{(k)})^T \nabla f(x^{(k)})}$$
$$+ \frac{1}{2}(x - x^{(k)})^T F(x^{(k)})(x - x^{k})$$
$$= q(x)$$

Please verify

$$\nabla q(x^*) = 0 \Rightarrow \nabla f(x^{(k)}) + F(x^k)(x^* - x^{(k)}) = 0$$
$$x^* = x^{(k)} - (F(x^{(k)}))^{-1} \nabla f(x^{(k)})$$

Next iterate for minimizer of f is

$$\boxed{x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1} \nabla f(x^{(k)})}$$

$=0$

$F > 0$

Algorithm

→ Solve $\qquad F(x^{(k)})\, d^{(k)} = -\nabla f(x^{(k)})$ for $d^{(k)}$ (Linear system)

→ Update $\qquad x^{(k+1)} = x^{(k)} + d^{(k)}$

Problem: If f is quadratic, and $F = F^T > 0$, the Newton's method converges in one step irrespective of the initial point $x^{(0)}$.

Pf: $\qquad f(x) = \frac{1}{2} x^T Q x - b^T x \qquad \Big| \quad \nabla f(x) = Qx - b, \quad F = Q$

Pf: $f(\underline{x}) = \frac{1}{2}\underline{x}^T Q \underline{x} - \underline{b}^T \underline{x}$

$\nabla f(\underline{x}) =$

$F = Q$

FONC gives

$Q\underline{x}^* - b = 0$

$\underline{x}^* = \underset{\sim}{Q^{-1}b}$

$$\underline{x}^{(1)} = \underline{x}^{(0)} - F(\underline{x}^0)^{-1} \nabla f(\underline{x}^0).$$

$$= \underline{x}^{(0)} - \underline{Q^{-1}\left[Q\underline{x}^0 - \underline{b}\right]}$$

$$x^{(1)} = \underline{x}^{(0)} - x^{(0)} + Q^{-1}\underline{b} = \underline{Q^{-1}\underline{b}} = \underline{x}^*.$$

Qns: (1) Solution of non-linear least-square problem using Newton's method

(2) Convergence of Newton's method