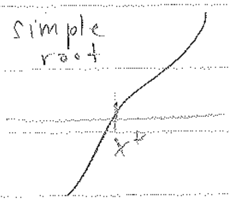
§6.3 Sensitivity & Conditioning

Solution to optimization problem inherently more sensitive than root finding

K ROOT FINDING



$K = \dfrac{1}{|f'(x^*)|}$

$\boxed{\text{back. err} \le K \text{ forw. err.}}$

$f(\hat{x}) \le \varepsilon \Rightarrow$

$|\hat{x} - x^*| \le K\varepsilon = \dfrac{\varepsilon}{|f'(x^*)|}$

K optimiz

$f(\hat{x}) = f(x^* + h) = f(x^*) + f'(x^*)h + \frac{1}{2} f''(x^*)h^2 + O(h^3)$

$f'(x^*) = 0 \quad (x^* \text{ is a minimum})$

$\Rightarrow f(x^* + h) \cong f(x^*) + \frac{1}{2} f''(x^*) h^2$

Then $|f(\hat{x}) - f(x^*)| \le \varepsilon \Rightarrow h \le \sqrt{\dfrac{2\varepsilon}{f''(x^*)}}$

if $\varepsilon = \varepsilon_{mach}$, error $\sqrt{\varepsilon_{mach}}$

Solution can be computed to only about $\frac{1}{2}$ digits

Min analogous to multiple root.

Keep in mind when selecting error tolerance

If $f''(x)$ available

Can directly solve $f'(x) = 0$

$K = \dfrac{1}{|f''(x^*)|}$

$|f'(\hat{x})| \le \varepsilon \Rightarrow |\hat{x} - x^*| \le \dfrac{\varepsilon}{|f''(x^*)|}$

$\boxed{\S 6.4}$ Optimization in 1D

1D $\begin{cases} \bullet \text{ important in own right} \\ \bullet \text{ subprob in higher D} \end{cases}$

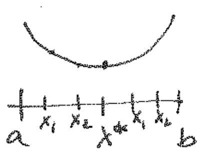
Bracketing

(root: sign change)

$f: \mathbb{R} \to \mathbb{R}$ unimodal on $[a,b]$

- $\exists\ x^* \in [a,b]$ s.t. $f(x^*) = \min$
- $x_1 < x_2$

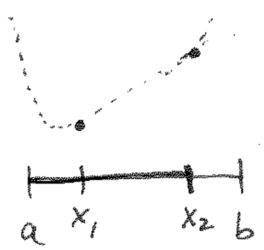
$$x_1 < x_2 < x^* \Rightarrow f(x_1) > f(x_2) \qquad x^* < x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$$



f ~~is~~ strictly decreasing $x \leq x^*$

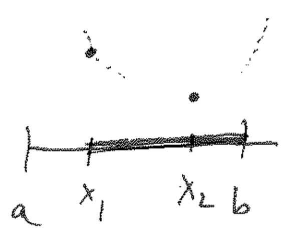
f strictly increasing $x \geq x^*$

$\boxed{\S 6.4.1}$ Golden Section Search.



$f(x_1) < f(x_2) \Rightarrow$ discard $(x_2, b]$

$[a, x_2]$

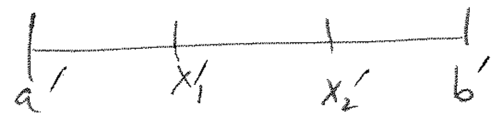


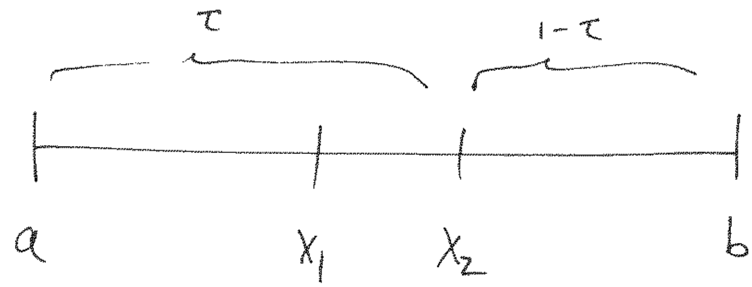
$f(x_1) > f(x_2) \Rightarrow$ discard $[a, x_1)$

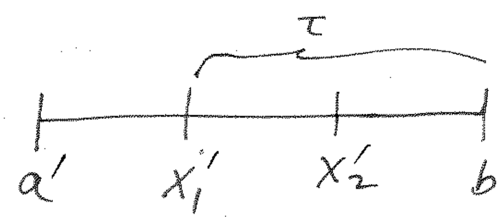
$[x_1, b]$

Next iteration: compute one new function evaluation.

Reduce length by same fraction each iteration







$$\boxed{\tau^2 = 1-\tau}$$

$$\tau^2+\tau-1 = 0$$

$$\boxed{\tau = \frac{-1+\sqrt{5}}{2}}$$

golden ratio

$\frac{\sqrt{5}-1}{2} \approx 0.618 \quad = \quad$ "golden ratio"

- length of new interval is $\tau \times$ prev. interval

$$\ell^{k+1} = \tau \ell^k$$

allows us to compute only 1 new function eval per iter.

- __safe__ for __unimodal functions__

- slow convergence

$r = 1$ linear

$C \cong .618$

- harder to find starting interval (than for root finding)

$[a, b]$

$x_1 = a + (1-\tau)(b-a)$

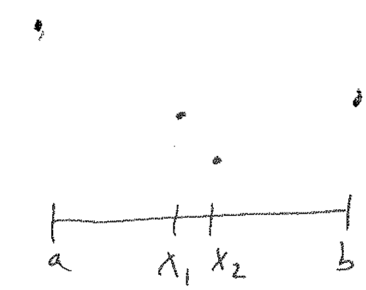
$x_2 = a + \tau(b-a)$

$f_1 = f(x_1)$

$f_2 = f(x_2)$

while $(b-a < \varepsilon)$

if $f_1 > f_2$, discard $[a, x_1)$



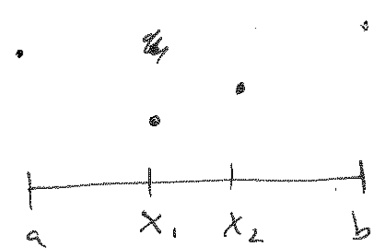
$a \leftarrow x_1$ still true

$x_1 \leftarrow x_2 = (a + (1-\tau)(b-a))$

$x_2 = a + \tau(b-a)$

$f_1 = f_2 \quad f_2 = f(x_2)$

else $f_1 < f_2$, discard $(x_2, b]$



$b \leftarrow x_2$ still true

$x_2 \leftarrow x_1 \quad (= a + \tau(b-a))$

$x_1 = a + (1-\tau)(b-a)$,

$f_2 = f_1$, $f_1 = f(x_1)$

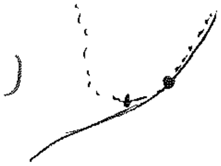
end

§ 6.4.3 Newton's Method

LECTURE 11

local quadratic approx. find min of quadratic.
(or find root of deriv.)

$$f(x+h) \approx f(x) + h f'(x) + \frac{h^2}{2} f''(x)$$



differentiate w.r.t. h, set to 0

$$f'(x) + h f''(x) = 0.$$

$$\Longrightarrow \boxed{h = -\frac{f'(x)}{f''(x)}}$$

x_0
for $k = 1, 2, \ldots$

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

Newton's Method on $f'(x) = 0$

end

- needs to be started near min
- may converge to <u>max</u> or <u>inflection pt</u>., or <u>fail</u>
- f' hard to find
- f'' also needed — worse.
- use secant type method to replace 2nd deriv. w. first derivatives
- could also replace 1st w/ function eval.

§ 6.4.4. Safeguarded Methods

hybrid methods.

e.g. successive parabolic interpolation + golden section search

§6.5 Unconstrained Optimization

Multi-dimensional.

§6.5.2 Steepest Descent Method.

$-\nabla f(x)$ direction of steepest descent (locally)

potent. useful direction to move
but step size ?

Define $\phi(\alpha) = f(\vec{x} + \alpha \vec{s})$ "line search" use a 1D solver.

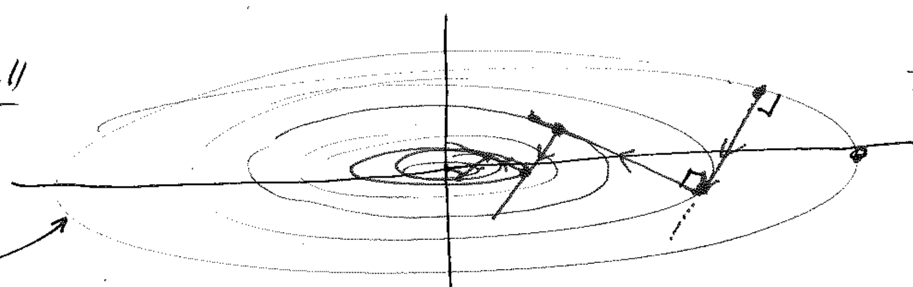
$\rightsquigarrow$ one-dimensional problem

$s = -\nabla f$ "steepest descent method"

x_0 = initial guess
for $k = 0, 1, 2, \ldots$
  $s_k = -\nabla f(x_k)$
  choose α_k to minimize $f(x_k + \alpha s_k)$ "line search"
  $x_{k+1} = x_k + \alpha_k s_k$
end

- always makes progress, but iterates can zigzag.
- linear conv, w/ factor arbitrarily close* to 1.

Example 6.11



$f(x) = 0.5x_1^2 + 2.5x_2^2$

$\nabla f = \begin{pmatrix} x_1 \\ 5x_2 \end{pmatrix}$

$\vec{x}_0 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$

1D opt. $\Rightarrow \alpha_0 = 1/3$

$\vec{x}_1 = \vec{x}_0 + \frac{1}{3}\vec{s}_0 = \begin{pmatrix} 5 \\ 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 3.333 \\ -.667 \end{pmatrix}$

- stop when $\|\nabla f\|$ small.

- contours where f = constant
- gradient @ $\vec{x}$ normal to level set

- min occurs when $\nabla f(\vec{x} + \alpha \vec{s}) \perp \vec{s}$