Today

First optimization algorithms
- Scalar
- Multivariable

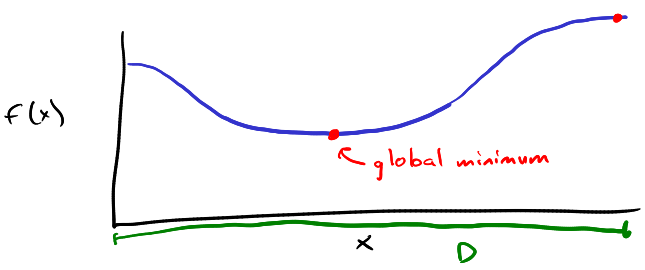
Announcements

Lab 3 due Thursday

Exam next Monday

Scalar Optimization





minimize $f(x)$
$x \in D$

Solve for $f'(x^*) = 0$ with ____ ↙ analytically for simple function

↑ favorite root finding nonlinear solver
incremental search, bisection, NR, secant

First order methods: use $f'(x)$ or $\nabla f(\vec{x})$ and (possibly) $f(x)$ or $f(\vec{x})$

Optimization and solving nonlinear eqns. are closely related

↷ Optimization ⇒ solving FONC (and SONC)

↳ solving $f(x)=0$ ⇒ minimize $f(x)^2$

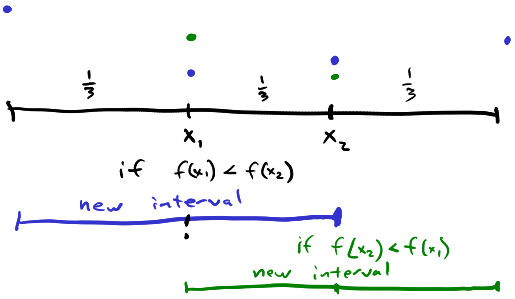
Zeroth order methods: use $f(x)$ or $f(\vec{x})$

Example:

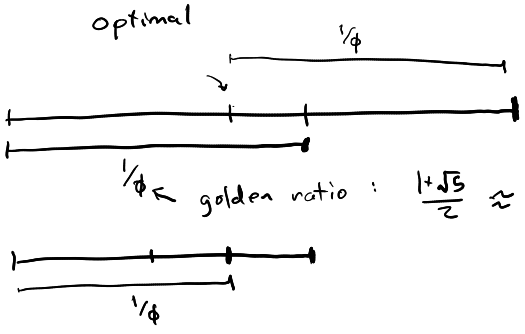
Golden-Section Search

input : function f, bracket $[x_l, x_u]$ containing a single minimum, tolerance ε_s
output: estimate x_{opt} of the minimizer

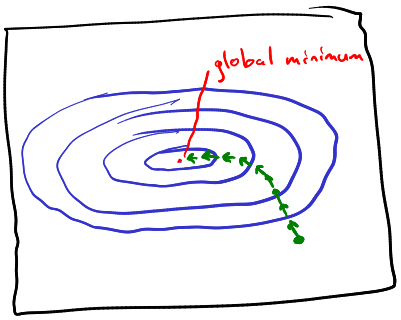
$R \leftarrow (\sqrt{5}-1)/2$ // golden ratio, ≈ 0.618
$d \leftarrow R(x_u - x_l)$
$x_1 \leftarrow x_l + d$; $x_2 \leftarrow x_u - d$
$f_1 \leftarrow f(x_1)$; $f_2 \leftarrow f(x_2)$
repeat
  $d \leftarrow Rd$
  // keep the subinterval with the smaller interior value
  if $f_1 < f_2$ **then**
    $x_l \leftarrow x_2$
    $x_2 \leftarrow x_1$; $f_2 \leftarrow f_1$
    $x_1 \leftarrow x_l + d$; $f_1 \leftarrow f(x_1)$
    $x_{\text{opt}} \leftarrow x_1$
  else
    $x_u \leftarrow x_1$
    $x_1 \leftarrow x_2$; $f_1 \leftarrow f_2$
    $x_2 \leftarrow x_u - d$; $f_2 \leftarrow f(x_2)$
    $x_{\text{opt}} \leftarrow x_2$
  end
  $\varepsilon_a \leftarrow (1-R)\left|\frac{x_u - x_l}{x_{\text{opt}}}\right| \cdot 100\%$
until $\varepsilon_a < \varepsilon_s$
return x_{opt}



What are optimal locations of x_1 and x_2



Multivariable Optimization



Gradient

$\nabla f(\vec{x})$: direction and magnitude of steepest ascent

$-\nabla f(\vec{x})$

↙ a.k.a. "steepest descent"

Gradient Descent (fixed step size)

input : starting point $\vec{x}_0$, step size α, tolerance E_s
output: $\vec{x}^*$, estimate of a local minimizer

$k \leftarrow 0$
while $\|\nabla f(\vec{x}_k)\| > E_s$ **do**
  $\vec{x}_{k+1} \leftarrow \vec{x}_k - \alpha \nabla f(\vec{x}_k)$ // step downhill
  $k \leftarrow k + 1$
end
return $\vec{x}^* = \vec{x}_k$

What should α be?

Unstable if $\alpha > \frac{2}{\lambda_{max}}$

Fixed step-size gradient descent is bad: either small steps or unstable.