One-Dimensional Optimization

CE 377K

January 29, 2015

REVIEW

Global minima, local minima, stationary points
Coercive function
One decision variable and multiple decision variables

ONE-DIMENSIONAL OPTIMIZATION

Assume we are solving a one-dimensional optimization problem of the form

$$\min_x f(x)$$

subject to $x \in [a, b]$

Why do we assume the interval $[a, b]$ is closed?

Since there are constraints, we need to change the definition of a local minimum slightly:

x^* is a local minimum if there is some interval (ℓ, h) containing x^* such that $f(x^*) \leq f(x)$ if $x \in (\ell, h) \cap [a, b]$.

A one-dimensional function is *unimodal* on $[a, b]$ if it is continuous and there is exactly one local minimum in $[a, b]$

(In this case, the local minimum must also be the global minimum.)

If f is unimodal, there are several *line search* techniques that can be used to solve the problem.

Since there are constraints $x \geq a$ and $x \leq b$, $f'(x)$ may not be zero at the optimal solution.

If f is differentiable, there are three possibilities for the optimal solution x^*:

$$\begin{aligned} \text{Case I} &: x \in (a, b) \text{ and } f'(x) = 0 \\ \text{Case II} &: x = a \text{ and } f'(x) \geq 0 \\ \text{Case III} &: x = b \text{ and } f'(x) \leq 0 \end{aligned}$$

A line search technique searches over the interval $[a, b]$ in search of a point satisfying one of these cases.

There are many different line search techniques that can be used; different techniques require different assumptions on f.

If f is twice differentiable, *Newton's method* is available.

If f is differentiable, the *bisection method* is available.

If f is merely continuous, the *golden section method* works.

We'll cover Newton's method and bisection in class; the homework will explore golden section.

NEWTON’S METHOD

Assume first that the optimum solution happens in the first case, where $f'(x) = 0$.

Then finding the optimum solution is as simple as finding the zero of f'.

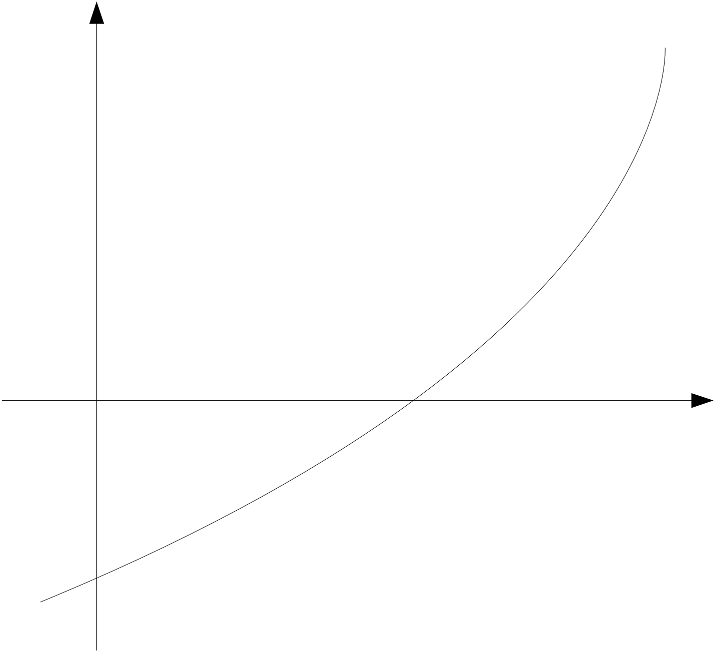
So, if the

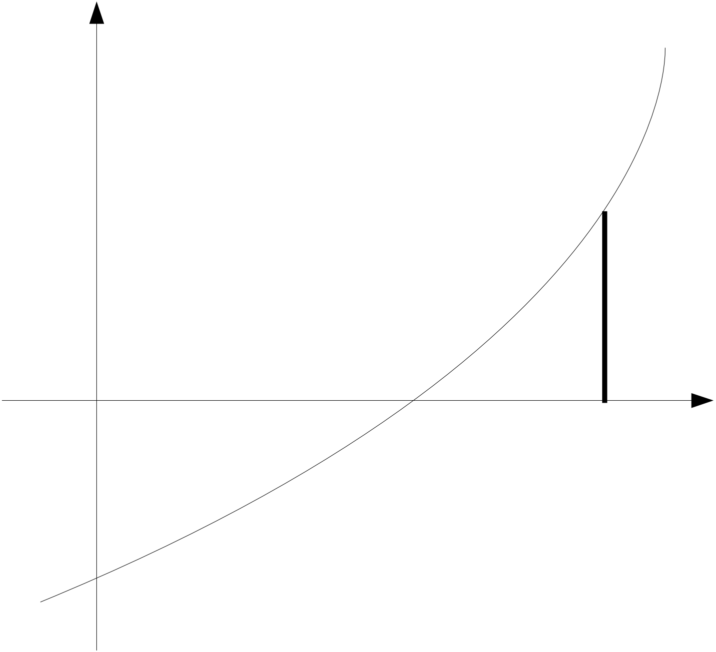
Remember the basic idea of Newton's method for an arbitrary function g.

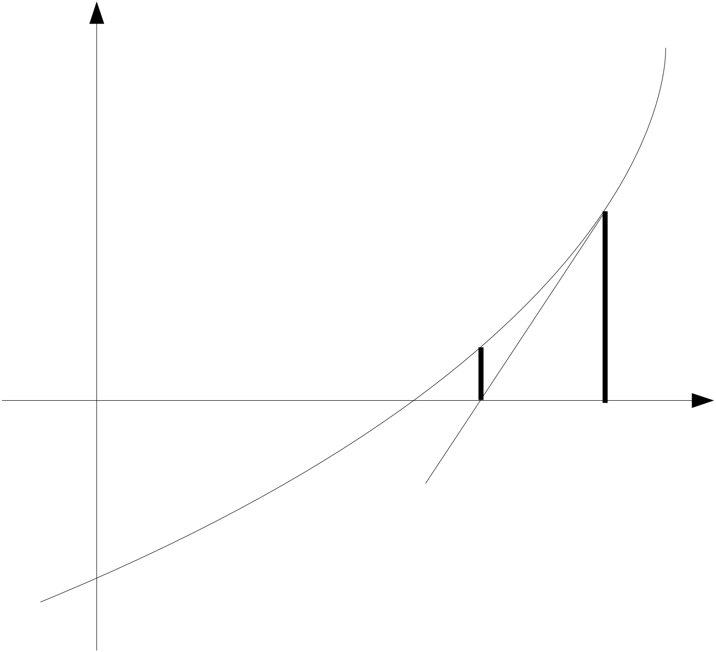
Starting with an initial guess x, approximate g with its linear approximation at that point: $g(y) \approx g(x) + g'(x)(y - x)$.

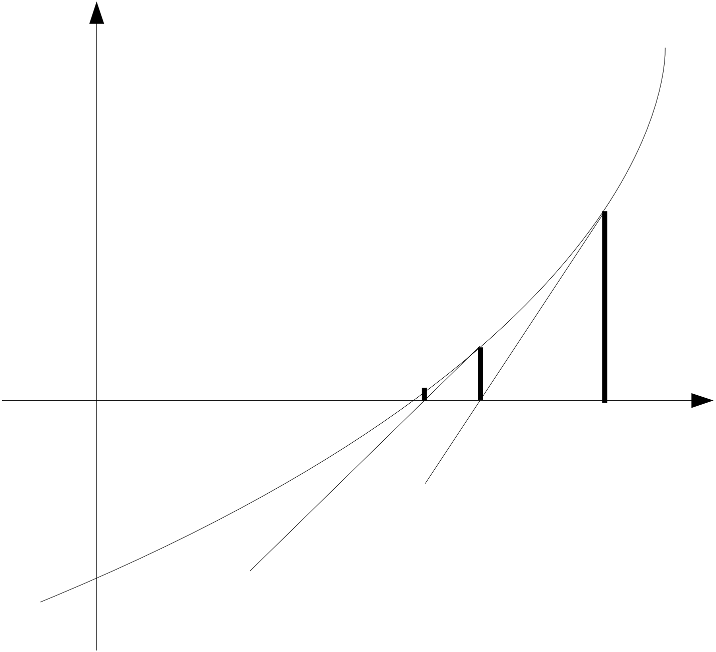
In this case, the zero happens at $x - g(x)/g'(x)$

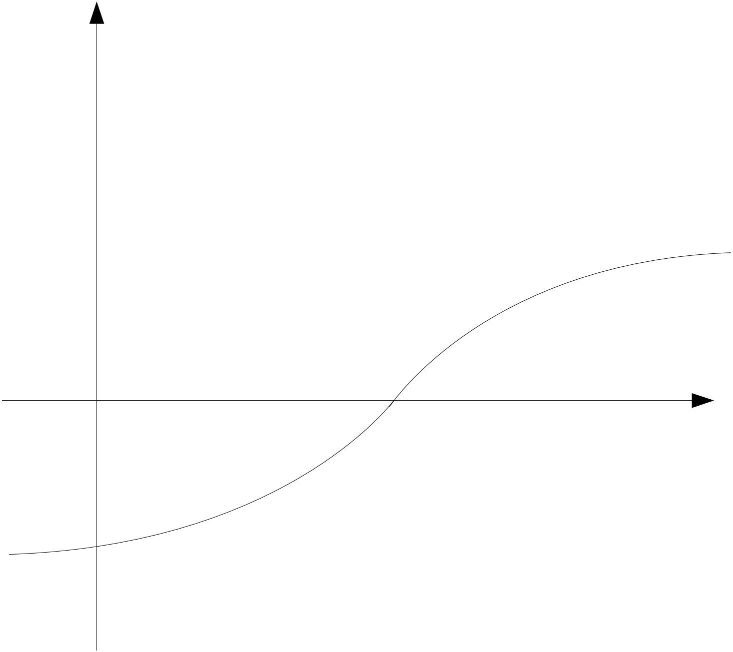
So, we update $x \leftarrow x - g(x)/g'(x)$ and repeat the process.

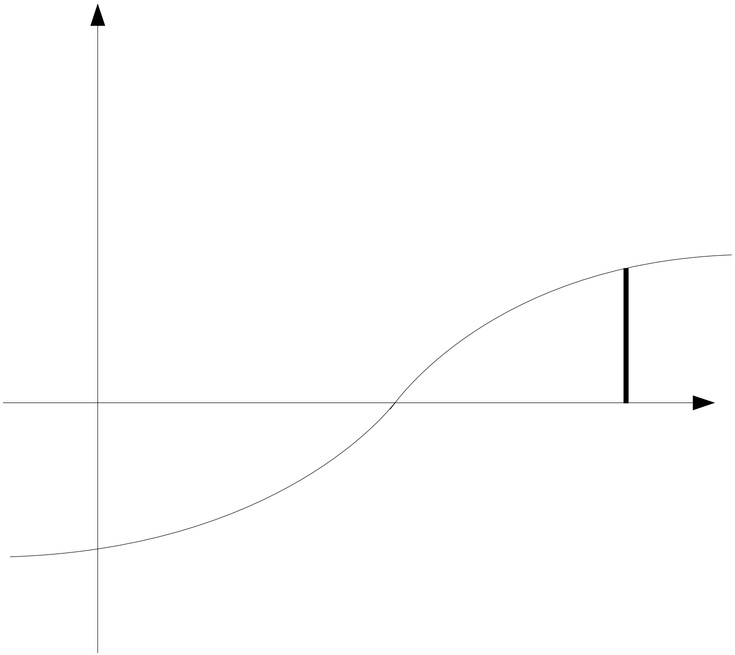


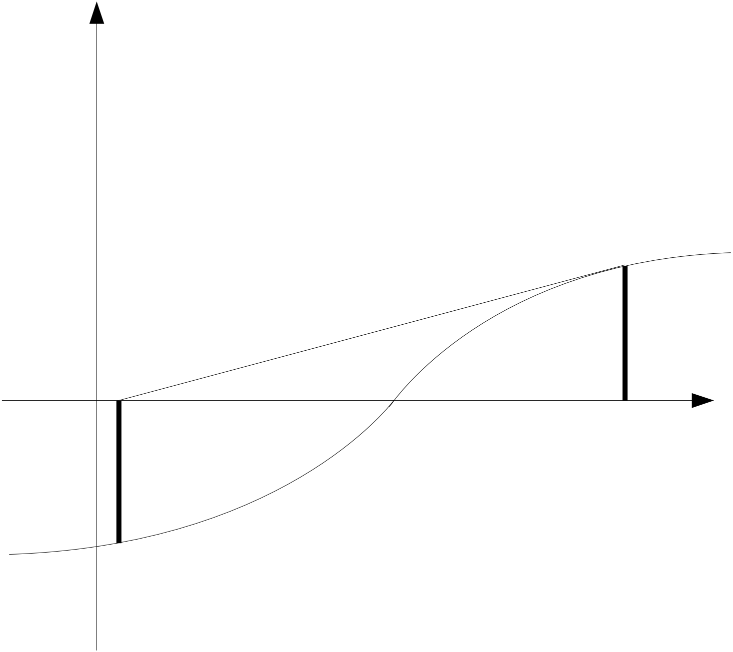


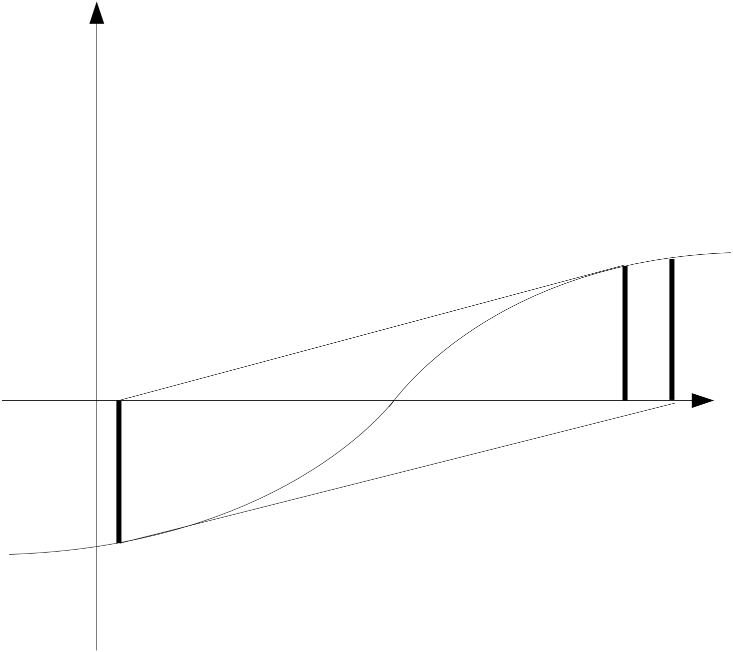


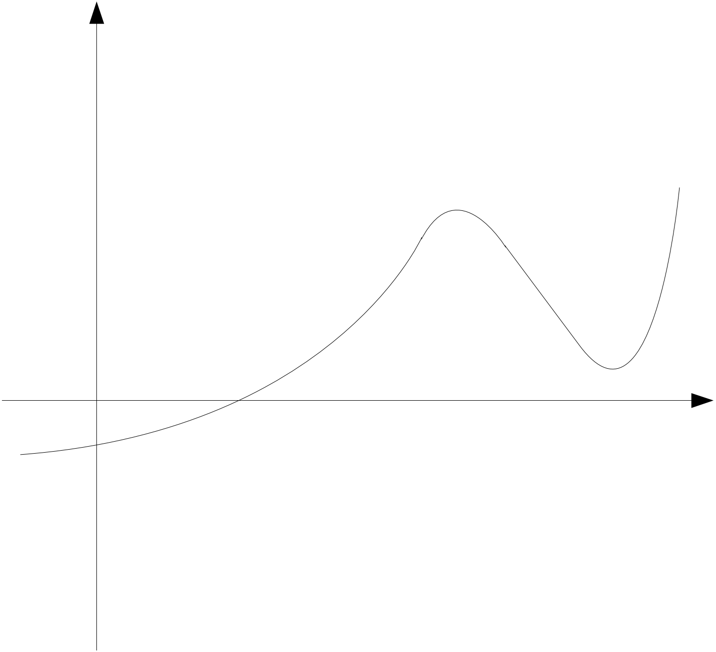


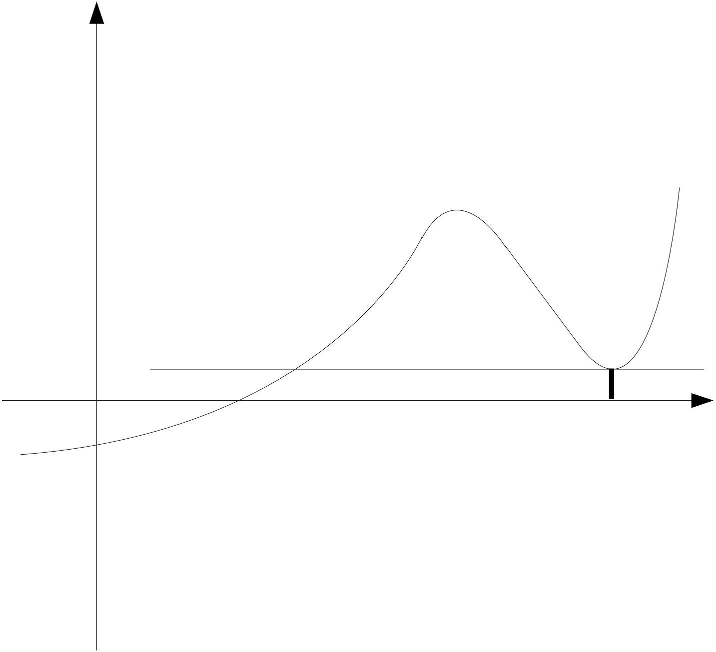












Newton's method will always work if we also assume that f is a convex function.

If f is twice differentiable, it is *convex* if $f''(x) \geq 0$ for all $x \in [a, b]$.

(There are more general definitions of convexity we will cover later in this course.)

What about the constraint $x \in [a, b]$?

Newton's method specialized for finding the minimum of a convex, twice differentiable function f:

1. Make an initial guess x
2. Calculate $f'(x)$ and $f''(x)$
3. If $f'(x) \approx 0$, stop.
4. Update $x \leftarrow x - f'(x)/f''(x)$
5. If $x < a$, set $x \leftarrow a$; if $x > b$ set $x \leftarrow b$.
6. Return to step 2.

If $f''(x) = 0$ at some point, move all the way to a if $f'(x) \geq 0$ or all the way to b if $f'(x) \leq 0$.

Example

Use Newton's method to find the minimum of $f(x) = (x-1)^4 + e^x$ on the interval $x \in [0, 3]$, performing seven iterations.

The first and second derivatives of $f(x)$ are

$$f'(x) = 4(x-1)^3 + e^x$$

and

$$f''(x) = 12(x-1)^2 + e^x$$

.

As an initial guess, choose the midpoint: $x = 1.5$. Here $f'(x) = 4.982$ and $f''(x) = 7.482$

The new value of x is $1.5 - 4.982/7.482 = 0.834$. Here $f'(x) = 2.285$ and $f''(x) = 2.633$

The new value of x is $0.834 - 2.285/2.633 = -0.034$. This is less than $a = 0$, so we set $x \leftarrow 0$.

And so on...

Iteration	x	$f'(x)$	$f''(x)$	$x - f'(x)/f''(x)$
1	1.5	4.982	7.482	0.834
2	0.834	2.285	2.633	−0.034
3	0	−3.000	13.000	0.231
4	0.231	−0.561	8.360	0.298
5	0.298	−0.0375	7.263	0.30304
6	0.30304	-2.058×10^{-4}	7.1829	0.3030725347
7	0.30307	-6.307×10^{-9}	7.1825	0.3030725355

Our discussion to date assumed the optimum solution is the first case. Does Newton's method still work in the other cases?

BISECTION METHOD

Another line search technique is the *bisection method*.

Advantages and disadvantages, relative to Newton's:

- (+) It only requires the function to be differentiable once, not twice.
- (+) It does not require calculation of a second derivative (so each iteration if faster).
- (-) Convergence usually requires more iterations.

Bisection works by narrowing the interval where the optimum solution must be found.

Initially, we only know that the optimum solution is somewhere in $[a, b]$. How can we narrow down the search?

Calculate the value of the derivative $f'(x)$ at the midpoint $x = (a + b)/2$.

What does the sign of $f'(x)$ tell us about where the optimum solution must be?

If $f'(x) > 0$, then the optimum solution must lie in the lower half of the interval $[a, (a+b)/2]$.

If $f'(x) < 0$, then the optimum solution must lie in the upper half of the interval $[(a+b)/2, b]$.

(What if $f'(x) = 0$?)

We can repeat this process over and over, halving the interval at each iteration.

Eventually, the interval is narrow enough and we can stop with the approximate optimum.

Algorithm

Step 0: Initialize. Set the iteration counter $k \leftarrow 1$, $a_1 \leftarrow a$, $b_1 \leftarrow b$.

Step 1: Evaluate midoint. Calculate the derivative of f at the middle of this interval, $d_k \leftarrow f'((a_k + b_k)/2)$

Step 2: Bisect. If $d_k > 0$, set $a_{k+1} \leftarrow a_k$, $b_{k+1} \leftarrow (a_k + b_k)/2$. Otherwise, set $a_{k+1} \leftarrow (a_k + b_k)/2$, $b_{k+1} = b$.

Step 3: Iterate. Increase the counter k by 1 and check the termination criterion. If $b_k - a_k < \epsilon$, then terminate; otherwise, return to step 1.

Example

Use the bisection method to find the minimum of $f(x) = (x-1)^4 + e^x$ on the interval $x \in [0, 3]$, performing seven iterations.

The derivative of $f(x)$ is

$$f'(x) = 4(x-1)^3 + e^x$$

.

Initially the interval is $[0, 3]$, and $f'(1.5) = 4.98 > 0$

The next interval is $[0, 1.5]$, and $f'(0.75) = 2.05 > 0$

The next interval is $[0, 0.75]$, and $f'(0.375) = 0.478 > 0$

The next interval is $[0, 0.375]$, and $f'(0.1875) = -0.939 < 0$

The next interval is $[0, 1875, 0.375]$, etc.

k	a_k	b_k	$(a_k + b_k)/2$	d_k
1	0	3	1.5	$4.98 > 0$
2	0	1.5	0.75	$2.05 > 0$
3	0	0.75	0.375	$0.478 > 0$
4	0	0.375	0.1875	$-0.939 < 0$
5	0.1875	0.375	0.28125	$-0.160 < 0$
6	0.28125	0.375	0.328125	$0.175 > 0$
7	0.28125	0.328125	0.3046875	$0.0116 > 0$

How can we be sure that the bisection method will always work?