ORF 307
Intro to NLP

Prof. John Mulvey

Spring, 2013

NLP

Find $x = (x_1, x_2, \ldots, x_n)$

$$\max \quad f(x)$$

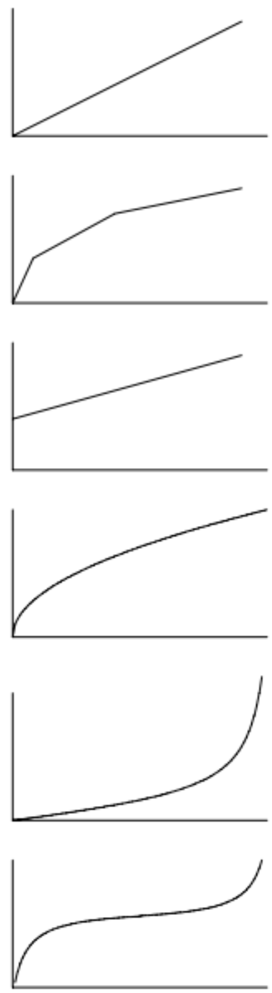
$$\text{s.t.}$$

$$g_i(x) \le b_i, \qquad \text{for } i = 1, 2, \ldots, m$$

$$x \ge 0$$

Note that this is an LP
if f(x) and all g_i(x) are linear.

Common cost functions



(a) Linear costs

(b) Piecewise linear costs

(c) Fixed + variable linear costs

(d) Concave nonlinear costs

(e) Convex nonlinear costs

(f) Combined nonlinear costs

Types of NLP

- There are many flavors of NLPs (several of these are subsets of others)

Name	f(x)	$g_i(x)$
Linear prog	linear	linear
Unconstrained opt	nonlinear/linear	none
Linearly constrainted prog	Nonlinear/linear	linear
Quadratic prog	Quadratic	linear
Convex (max)	Concave	convex
Convex (min)	Convex	convex
Non-convex	Anything else	

- There are other types: fractional, geometric, conic, semi-definite,……

Convex functions (1 var)

<u>Real analysis definition:</u>

A function f: $\mathbb{R} \to \mathbb{R}$ defined on (a,b) is said to be <u>convex</u> if it lies below it's support.

Mathematically:

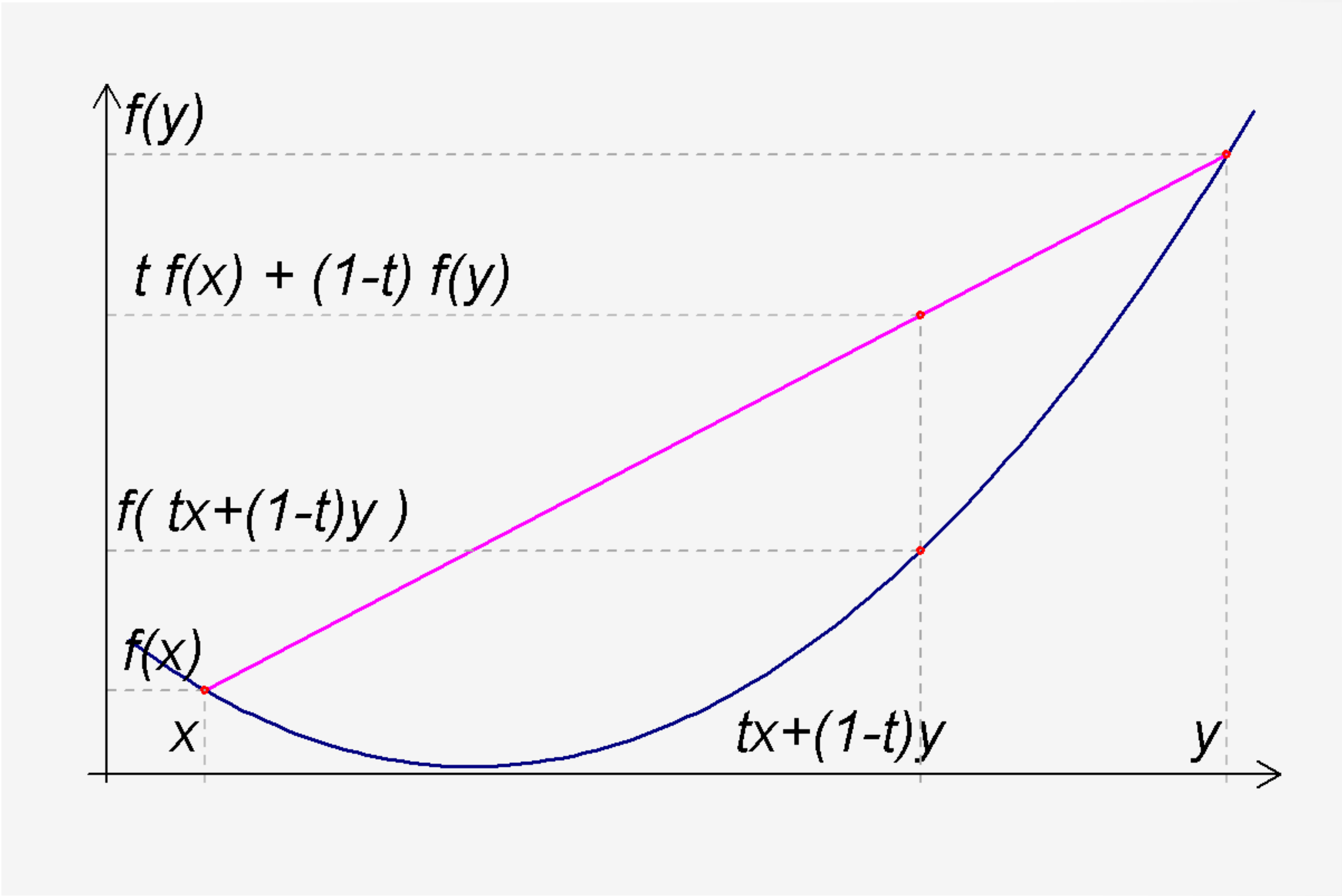
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

For all $x \,\epsilon (a,b)$, y $\epsilon (a,b)$ and $t \,\epsilon\, (0,1)$.

Definition: A function f is said to be <u>concave</u> if -f is convex.

<u>Theorem</u>

If f is twice differentiable, then f is convex on the real line <u>if and only if</u> $f''(x) \geq 0 \; \forall x$



f(y)
t f(x) + (1-t) f(y)
f(tx+(1-t)y)
f(x)
x
tx+(1-t)y
y

Convex functions (multi var)

Real analysis definition:

A function $f\colon\ \mathbb{R}^n \to \mathbb{R}$ defined on a convex set S is said to be convex if it lies below it's support.

Mathematically:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Same as before
But x, y are vectors

For all $x \in S$, y $\in S$ and $t \in (0,1)$.

Definition: A function f is said to be concave if -f is convex.

Theorem

If f is twice differentiable, then f is convex on the real line if and only if its Hessian is positive semi-definite.

Hessian matrix

$$H(f) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

A matrix H is said to be <u>positive-semidefinite</u> if $x^T H x \geq 0 \;\; \forall x \in \mathbb{R}^n$.

2-variable function

- $H(f) = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial y \partial x} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

- $H(f)$ is positive-semidefinite (and thus f is convex) if <u>all</u> of the following are true:
  - $a \geq 0$
  - $c \geq 0$
  - $ac - b^2 \geq 0$

Quadratic functions

- $f(\vec{x})$ is quadratic if there exists a symmetric matrix Q such that:
    - $f(\vec{x}) = x^T Q x$
- Example: $f(x_1, x_2) = x_1^2 - 4x_1x_2 + 3x_2^2$
- Here $Q = \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix}$
- Not quadratic:
    - $f(x) = x^3, f(x_1, x_2) = x_1^2 x_2, f(x) = x^{1.5}$

Hessian for quadratic functions

- In this case, the Hessian only contains constants.
- i.e.

$$H(f) = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

For some $a, b, c \in \mathbb{R}$

Examples of concave and convex functions

Convex	Concave
$ax + b$	$ax + b$
e^x	$\log(x)$
$\|x\|$	$\sqrt{x}$
x^3 when $x \geq 0$	x^3 when $x \leq 0$
x^k when k is even	-x^k when k is even

<u>Properties</u>

If f, g are convex functions, then the following functions are convex:

1. $h(x) = f(x) + g(x)$
2. $h(x) = \max(f(x), g(x))$
3. $h(x) = f(ax + b)$
4. If g is also nondecreasing, then $h(x) = g(f(x))$ is convex.
   For example if $f(x)$ is convex then $e^{f(x)}$ is convex.

Solving NLPs

- There are many algorithms to solve NLPs. (see IOR sec. 12.9)
- We will not cover these in detail.
- We will mainly use AMPL and Excel.
- Some questions can be solved by hand using the famous KKT conditions which will be introduced next class.
- It is important to know for which problems we will get a **global** optimum.

Global vs. local optima

- Algorithms can "get stuck" on local optima or at saddle points.

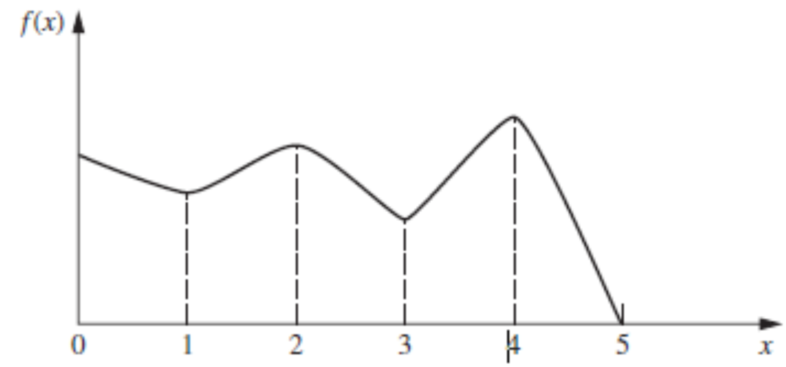


FIGURE 12.8
A function with several local maxima ($x = 0, 2, 4$), but only $x = 4$ is a global maximum.

A local maximum is a global maximum when the problem is a convex programming problem.

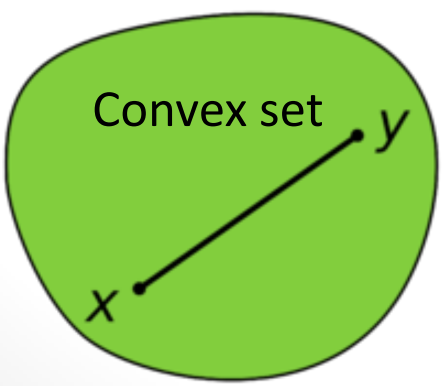
Note that LPs are convex programs

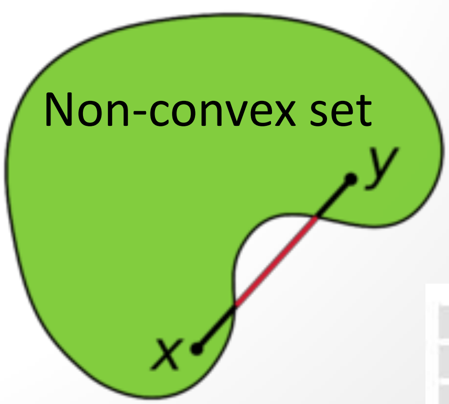
Convex Sets

- A convex set is a set of points that contains the entire line segment between its points.
- Formally:

S is said to be a <u>convex set</u> if $\forall x, y \in S$ and $\forall t \in [0,1]$:

$$(1-t)x + ty \in S$$





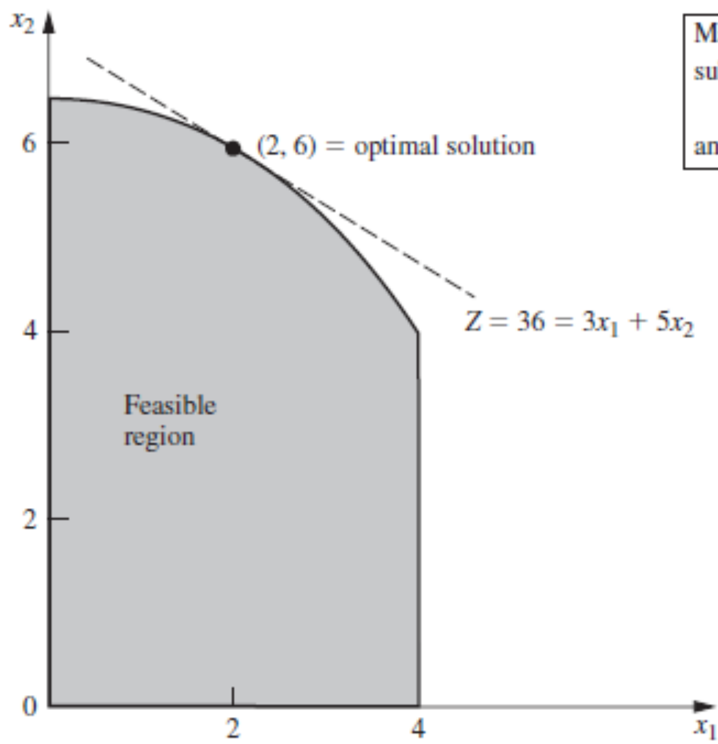
Constraint set

- Each inequality with a convex function describes a convex set.
- The feasible region is the intersection of all regions described by the constraints
- Fact: the intersection of convex sets is a convex set.
- Thus the feasible region is a convex set when the g_i's are convex.

The Objective

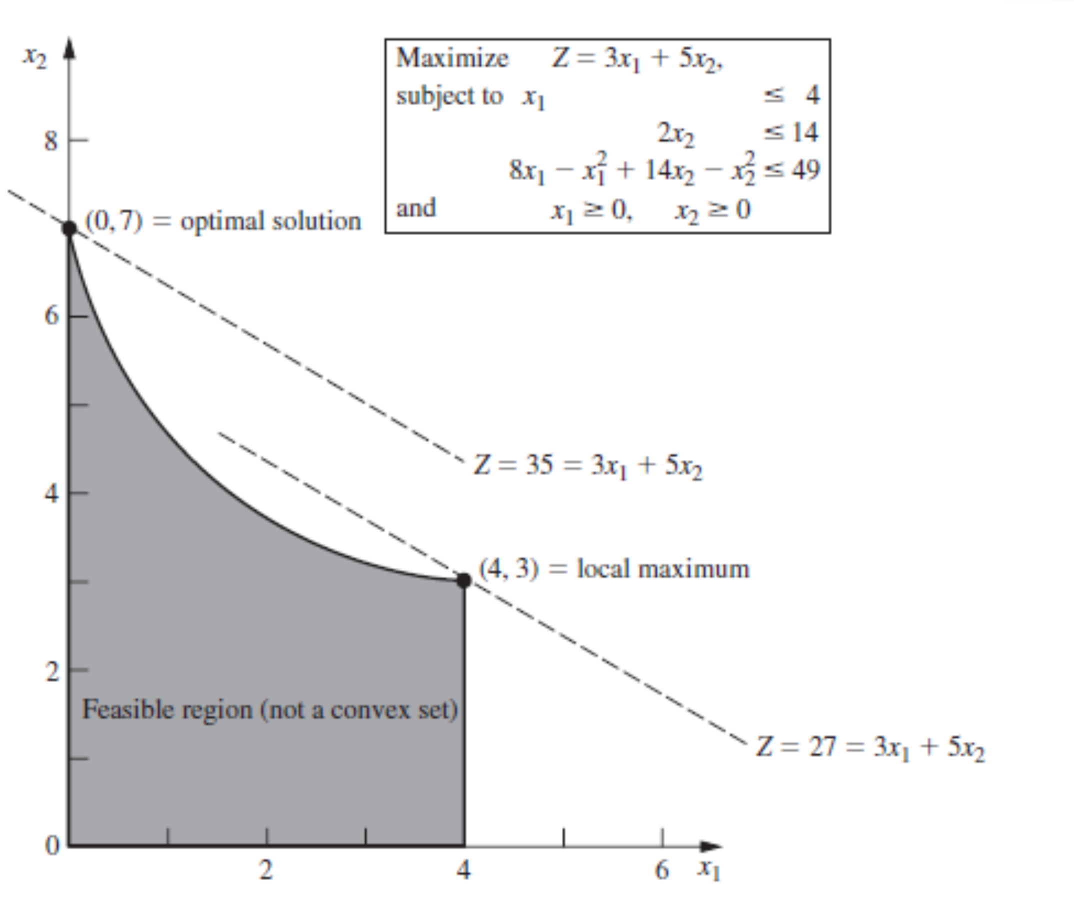
- A full proof is very technical.
- Given some regularity conditions, a concave function's local maximum is always a global maximum on a convex set.

Example with linear obj and convex feasible region



$$
\begin{array}{ll}
\text{Maximize} & Z = 3x_1 + 5x_2, \\
\text{subject to} & x_1 \leq 4 \\
 & 9x_1^2 + 5x_2^2 \leq 216 \\
\text{and} & x_1 \geq 0, \quad x_2 \geq 0
\end{array}
$$

Example with non-convex region



Example of a non-concave objective

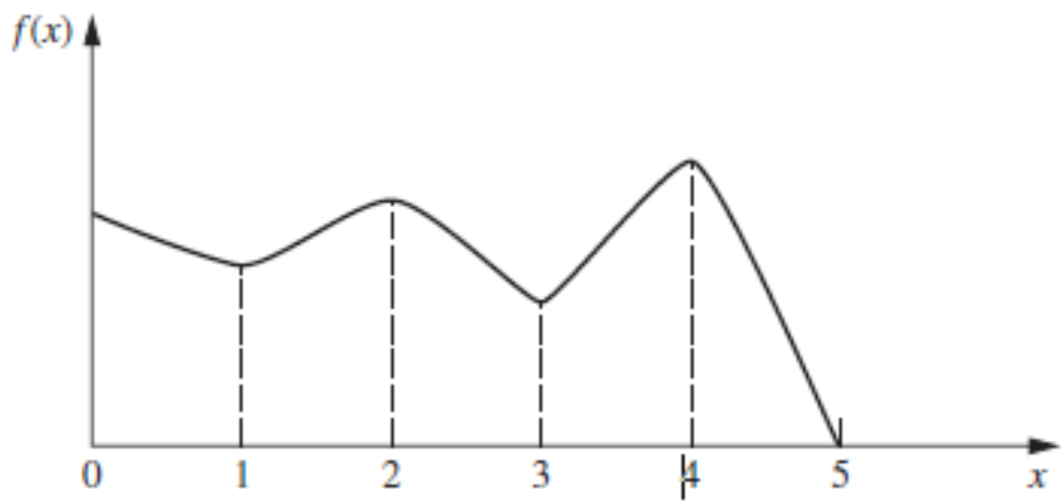


FIGURE 12.8
A function with several local maxima ($x = 0, 2, 4$), but only $x = 4$ is a global maximum.

Non-convex programming

- Most non-convex programs are hard to solve.
- Programs like excel can still solve these.
- However, solutions returned by these algorithms might only be locally optimal.
- One should use the multi-start option or use an algorithm with randomization.