ISE 426
Optimization models and applications

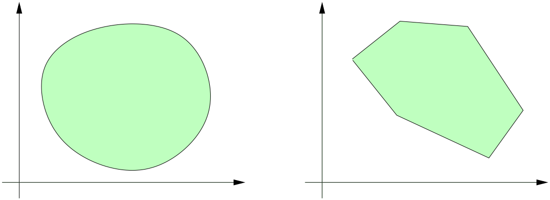
Lecture 2 — September 2, 2014

Convexity; Relaxations; Lower and upper bounds.

- Winston, chapter 1, **or**
- Winston & Venkataramanan, chapter 1, **or**
- Hillier & Lieberman, chapter 2.

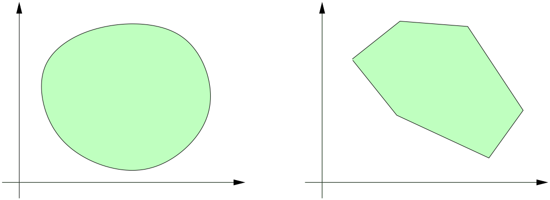
Convexity

Convex sets



Def.: A set $S \subseteq \mathbb{R}^n$ is convex if any two points x' and x'' of S are joined by a segment **entirely** contained in S:

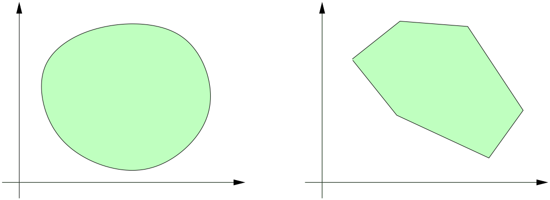
Convex sets



Def.: A set $S \subseteq \mathbb{R}^n$ is convex if any two points x' and x'' of S are joined by a segment **entirely** contained in S:

$$\forall x', x'' \in S, \alpha \in [0,1], \quad \alpha x' + (1-\alpha)x'' \in S$$

Convex sets

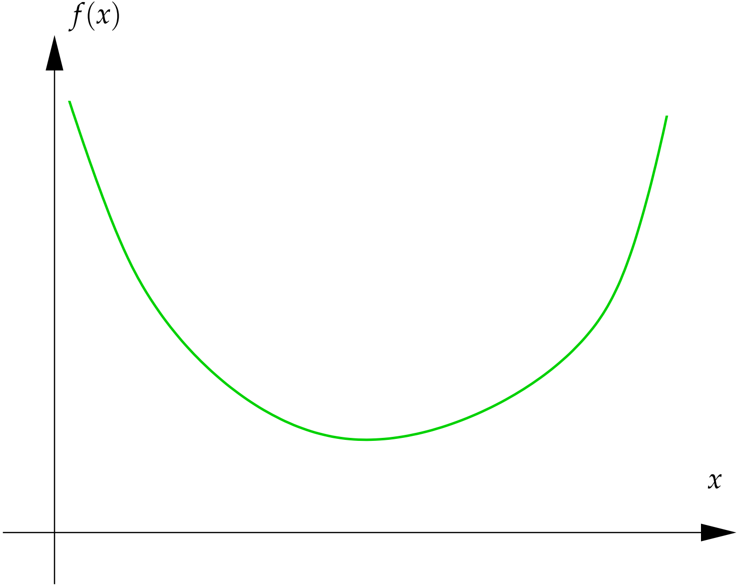


Def.: A set $S \subseteq \mathbb{R}^n$ is convex if any two points x' and x'' of S are joined by a segment **entirely** contained in S:

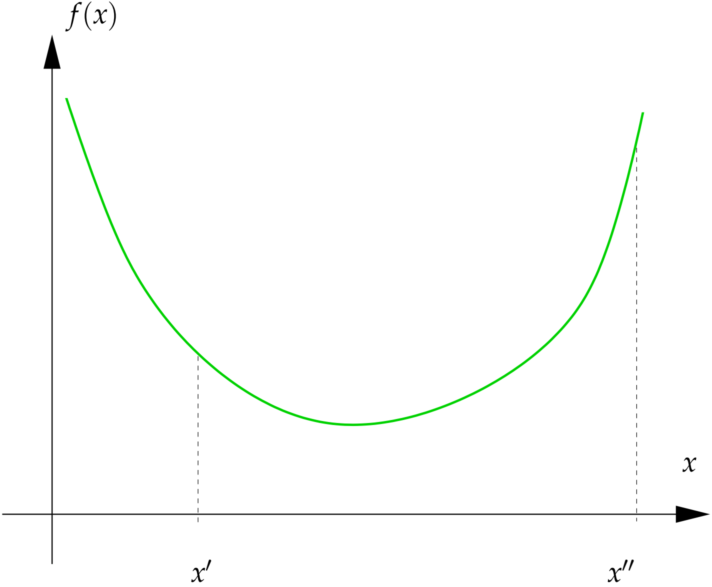
$$\forall x', x'' \in S, \alpha \in [0,1], \quad \alpha x' + (1-\alpha)x'' \in S$$

The intersection of two convex sets is convex.

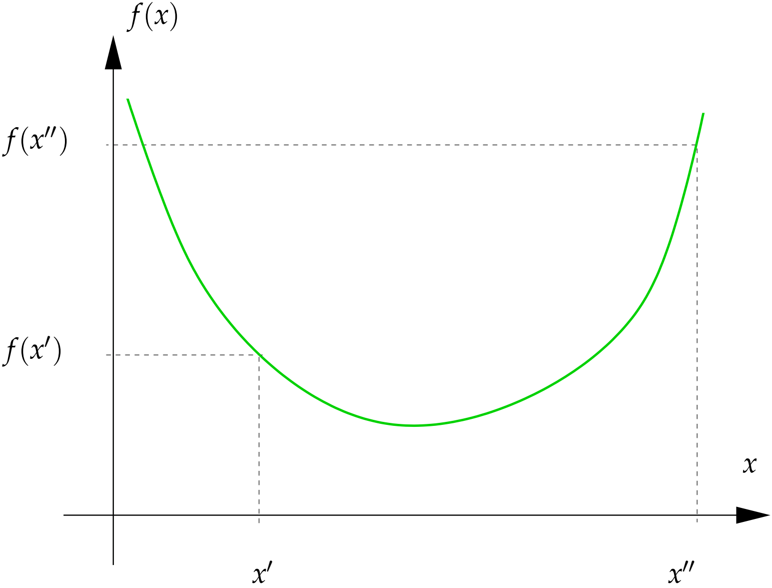
Definition



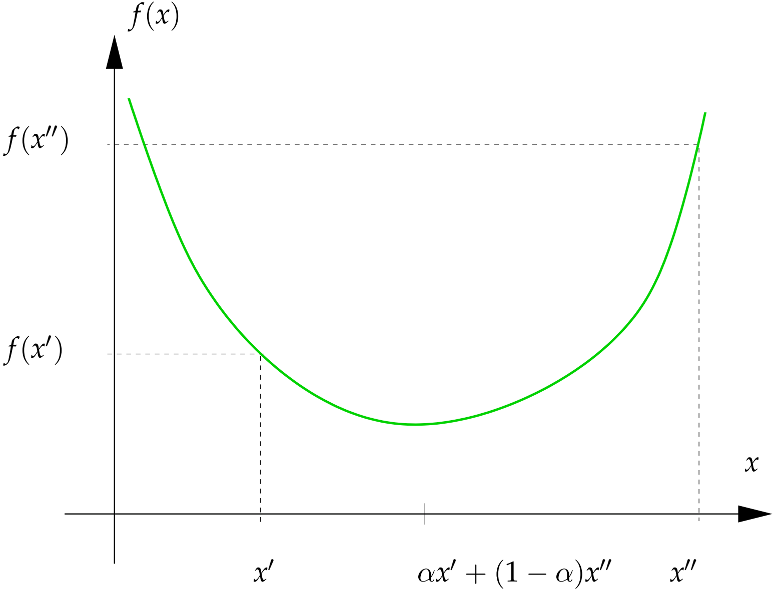
Definition



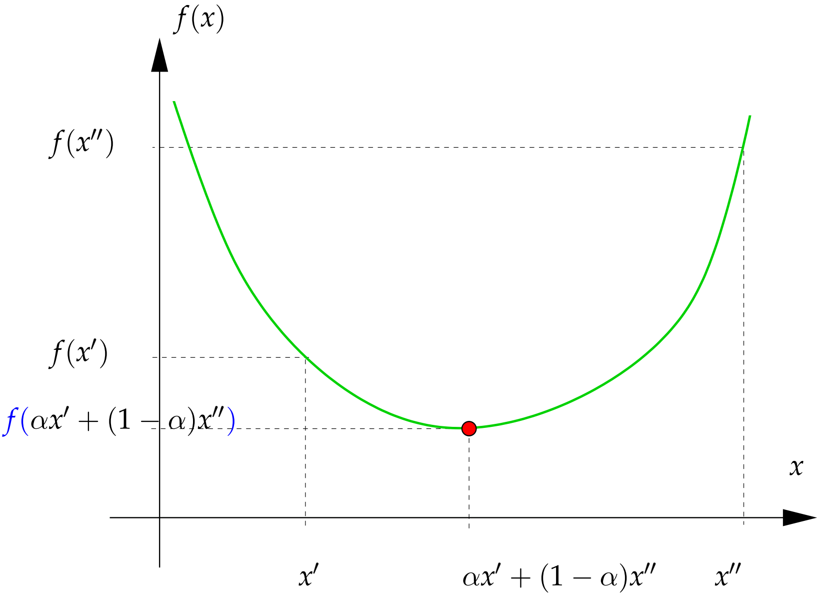
Definition



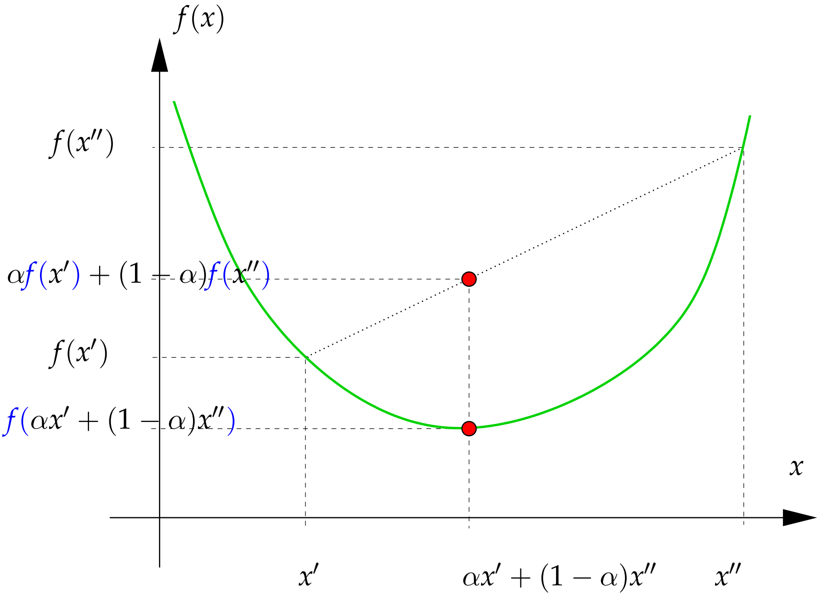
Definition



Definition



Definition


$f(x)$
$f(x'')$
$\alpha f(x') + (1-\alpha) f(x'')$
$f(x')$
$f(\alpha x' + (1-\alpha)x'')$
x
x'
$\alpha x' + (1-\alpha)x''$
x''

Convex functions

Def.: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if, for any two points x' and $x'' \in \mathbb{R}^n$ and for any $\alpha \in [0, 1]$

$$f(\alpha x' + (1-\alpha)x'') \quad \leq \quad \alpha f(x') + (1-\alpha) f(x'')$$

Convex functions

Def.: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if, for any two points x' and $x'' \in \mathbb{R}^n$ and for any $\alpha \in [0, 1]$

$$f(\alpha x' + (1-\alpha)x'') \quad \leq \quad \alpha f(x') + (1-\alpha) f(x'')$$

- The sum of convex functions is a convex function

Convex functions

Def.: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if, for any two points x' and $x'' \in \mathbb{R}^n$ and for any $\alpha \in [0, 1]$

$$f(\alpha x' + (1 - \alpha)x'') \quad \leq \quad \alpha f(x') + (1 - \alpha)f(x'')$$

- The sum of convex functions is a convex function
- Multiplying a convex function by a positive scalar gives a convex function

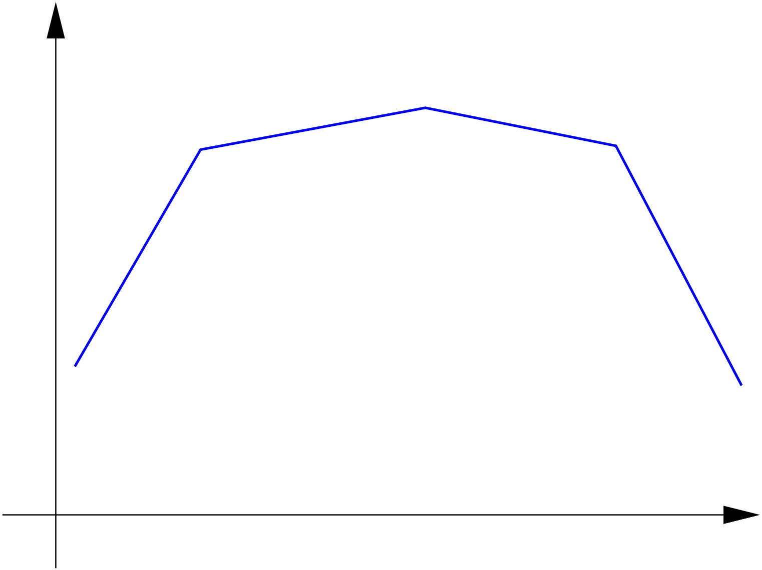
Convex functions

Def.: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if, for any two points x' and $x'' \in \mathbb{R}^n$ and for any $\alpha \in [0, 1]$

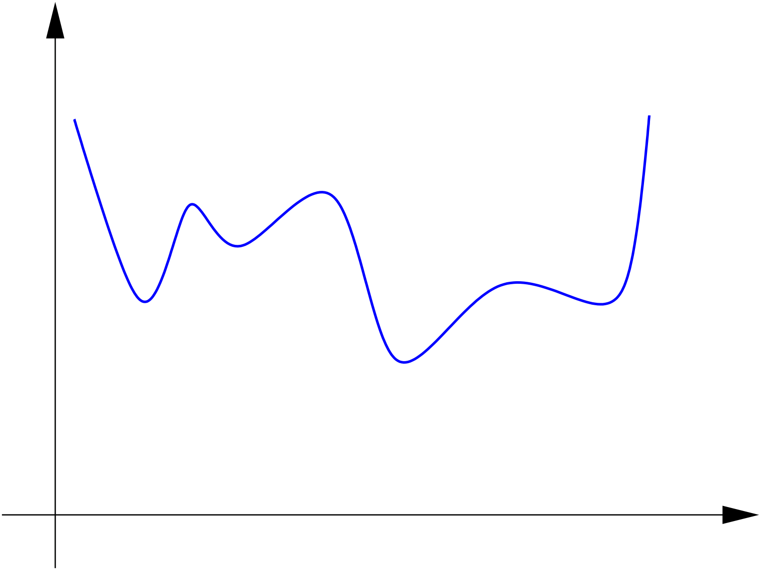
$$f(\alpha x' + (1 - \alpha)x'') \quad \leq \quad \alpha f(x') + (1 - \alpha)f(x'')$$

- The sum of convex functions is a convex function
- Multiplying a convex function by a positive scalar gives a convex function
- **linear** functions $\sum_{i=1}^{k} a_i x_i$ are convex, irrespective of the sign of a_i's.

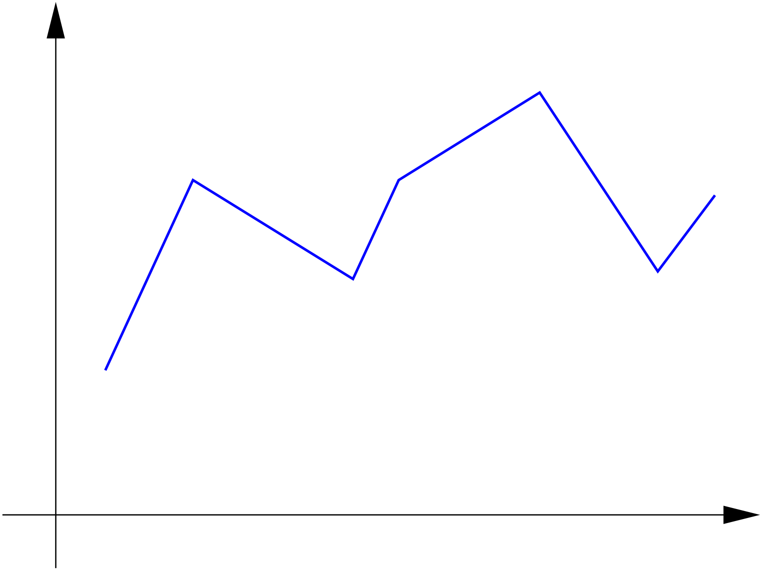
Definition



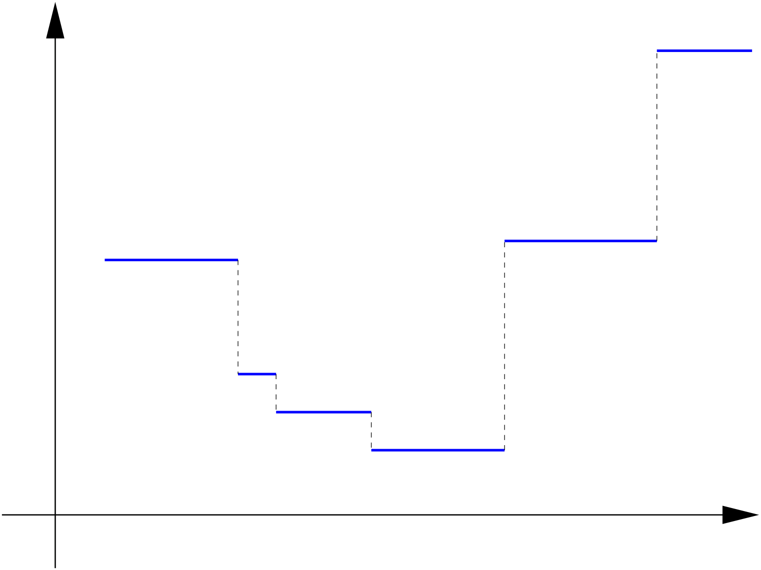
Definition



Definition



Definition



Local and global optima

A vector $x^l \in \mathbb{R}^n$ is a local optimum if

- there is a *neighbourhood*[1] N of x^l with no better point than x^l:

$$\forall x \in N, f_0(x) \geq f_0(x^l)$$

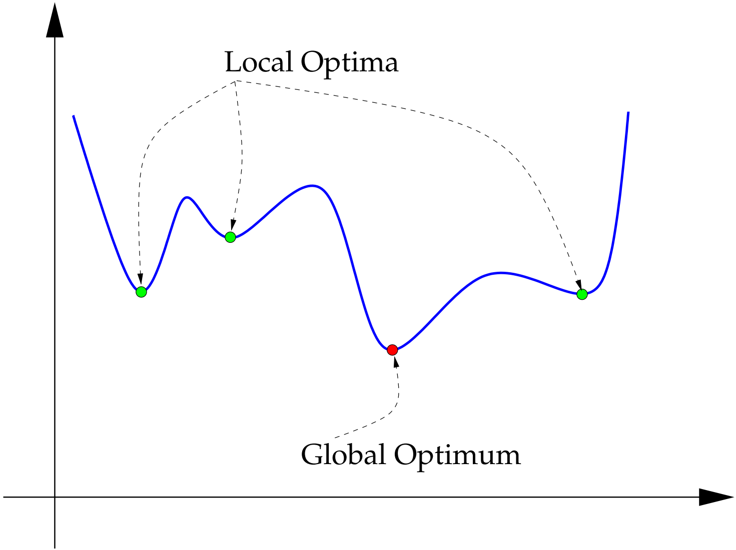
A vector $x^g \in \mathbb{R}^n$ is a global optimum if

- x^g
- there is no x better than x^g, i.e.,

$$f_0(x) \geq f_0(x^g) \qquad \forall x$$

[1] A *neighbourhood* of x^l can be defined as $N = \{x : ||x - x^l||_2 \leq \epsilon\}$ for some ϵ.

Local optima, global optima



Positive (Semi)Definite Matrices

A square $n \times n$ matrix A is **Positive Definite** (PD) (denoted with $A \succ 0$) if, for any n-vector $x \neq 0$, the following holds:

$$x^\top A x \geq 0$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j =$$

Positive (Semi)Definite Matrices

A square $n \times n$ matrix A is **Positive Definite** (PD) (denoted with $A \succ 0$) if, for any n-vector $x \neq 0$, the following holds:

$$x^\top A x \geq 0$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j =$$

$$\begin{array}{llll}
a_{11}x_1^2 & +a_{12}x_1x_2 & +\dots & +a_{1n}x_1x_n + \\
a_{21}x_2x_1 & +a_{22}x_2^2 & +\dots & +a_{2n}x_2x_n + \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1}x_nx_1 & +a_{n2}x_nx_1 & +\dots & +a_{nn}x_n^2
\end{array}$$

$$> 0$$

Some linear algebra

- A **minor** of a $m \times n$ matrix A is the determinant of a (square) submatrix of A obtained by removing some rows and columns of A.
- For a square matrix B, a **principal minor** of B is obtained by removing the same row and column indices from B
- For a square matrix B, a **leading principal minor** of B is a principal minor obtained by removing the last k rows and (the same) columns from B

For example, $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 7 \\ 3 & 6 & 8 & 9 \\ 4 & 7 & 9 & 0 \end{pmatrix}$.

Leading principal minors: (1), $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 8 \end{pmatrix}$

Positive (Semi)Definite Hessian Matrices

A symmetric matrix $A \in R^{n \times n}$ is **Positive Semidefinite** (PSD) (denote it with $A \succeq 0$) if all principal minors of A are nonnegative.

A symmetric matrix $A \in R^{n \times n}$ is **Positive Semidefinite** (PSD) if and only if $A = BB^\top$ for some $B \in R^{n \times n}$.

A symmetric matrix $A \in R^{n \times n}$ is **Positive Semidefinite** (PSD) if for all $i = 1, \ldots, n$ $a_{ii} \geq \sum_{j=1, j \neq i}^{n} |a_{ij}|$.

A symmetric matrix $A \in R^{n \times n}$ is **Positive Semidefinite** (PSD) if all its eigenvalues are nonnegative.

Positive (Semi)Definite Hessian Matrices

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, consider its **Hessian**:

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** on set $\Omega \in R^n$ if and only if the Hessian $H_f(x) \succeq 0$ for all $x \in \Omega$.

Examples

- The function $f(x) = x$ is convex
- The function $f(x_1, x_2) = x_1 + x_2$ is convex
- The function $f(x_1, x_2) = x_1^2 + x_2$ is convex
- The function $f(x_1, x_2) = 5x_1^2 + 3x_2^2$ is convex
- The function $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2$ is convex

Examples

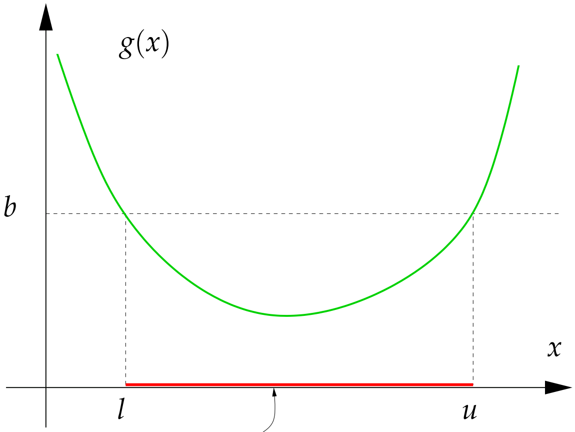
- The function $f(x) = x$ is convex
- The function $f(x_1, x_2) = x_1 + x_2$ is convex
- The function $f(x_1, x_2) = x_1^2 + x_2$ is convex
- The function $f(x_1, x_2) = 5x_1^2 + 3x_2^2$ is convex
- The function $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2$ is convex

- The function $f(x_1, x_2) = x_1^2 + x_2^2 + 5x_1x_2$ is nonconvex
- The function $f(x_1, x_2) = x_1^2 - x_2^2$ is nonconvex
- The function $f(x_1, x_2) = x_1x_2$ is nonconvex
- The function $f(x) = \sin x$, for $x \in [0, 2\pi]$ is nonconvex
- The function $f(x) = -x^2$ is nonconvex

Convex constraints

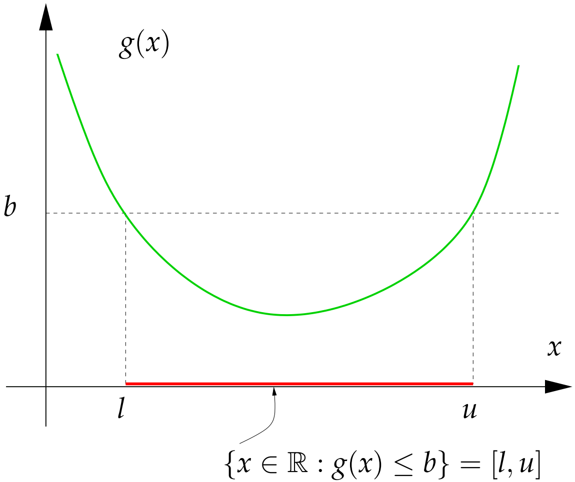
- A constraint $g(x) \leq b$, with $g : \mathbb{R}^n \to \mathbb{R}$, defines a subset S of $\mathbb{R}^n$, that is,
$$S = \{x \in \mathbb{R}^n : g(x) \leq b\}$$
- The constraint $g(x) \leq b$ is convex if the set S is convex.
- If the *function* $g(x)$ is convex, the *constraint* $g(x) \leq b$ is convex.

Convex constraints



$$\{x \in \mathbb{R} : g(x) \leq b\} = [l, u]$$

Convex constraints



Note: if the *function* $g(x)$ is convex, the *constraint* $g(x) \geq b$ may be nonconvex!

Convex and concave functions

Def.: A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $-f(x)$ is convex.

- concave functions are useful with **maximization** problems:

$$\begin{aligned} \max \quad & f(x) \\ & g_i(x) \leq 0 \quad \forall i = 1, 2 \ldots, m \\ & h_i(x) = 0 \quad \forall i = m+1, m+2 \ldots, m+q \\ & x \in \mathbb{R}^n \end{aligned}$$

  is a convex problem if $f(x)$ is concave, all $g_i(x)$ are convex, and all $h_i(x)$ are affine.

- concave functions are also useful for $\geq$ constraints:

$$g_i(x) \geq 0$$

  is a convex constraint if $g_i(x)$ is concave.

For what convex $g(x)$ the *constraint* $g(x) \geq b$ is always convex?

For what convex $g(x)$ the *constraint* $g(x) \geq b$ is always convex?

- linear constraints $\sum_{i=1}^{k} a_i x_i \left\{ \begin{array}{c} \leq \\ = \\ \geq \end{array} \right\} b$ are convex

The general optimization problem

Consider a vector $x \in \mathbb{R}^n$ of variables.
An optimization problem can be expressed as:

$$
\begin{aligned}
\mathbf{P}: \qquad \text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & f_1(x) \leq b_1 \\
& f_2(x) \leq b_2 \\
& \vdots \\
& f_m(x) \leq b_m
\end{aligned}
$$

Feasible solutions, local and global optima

Define $F = \{x \in \mathbb{R}^n : f_1(x) \leq b_1, f_2(x) \leq b_2, \ldots, f_m(x) \leq b_m\}$, that is, F is the feasible set of an optimization problem.
All points $x \in F$ are called feasible solutions.

A vector $x^l \in \mathbb{R}^n$ is a local optimum if

- $x^l \in F$
- there is a *neighbourhood*[2] N of x^l with no better point than x^l:

$$\forall x \in N \cap F, f_0(x) \geq f_0(x^l)$$

A vector $x^g \in \mathbb{R}^n$ is a global optimum if

- $x^g \in F$
- there is no $x \in F$ better than x^g, i.e.,

$$f_0(x) \geq f_0(x^g) \qquad \forall x \in F$$

[2] A *neighbourhood* of x^l can be defined as $N = \{x : ||x - x^l||_2 \leq \epsilon\}$ for some ϵ.

Convex problems

Def.: An optimization problem is convex if

- the objective function is convex
- **all** constraints are convex

Convex problems

Def.: An optimization problem is convex if

- the objective function is convex
- **all** constraints are convex

Convex optimization problems are *easy*:

> **If a problem P is convex, a local optimum $x^\star$ of P is also a global optimum of P.**

(Hint) When modeling an optimization problem, it would be good if we found a convex problem.

Examples of convex problems

$$
\mathbf{P}: \quad \begin{array}{rl} \min & x_1^2 + 2x_2^2 \\ \text{s.t.} & x_1^2 + x_2^2 \leq 1 \\ & 0 \leq x_1 \leq 2 \\ & 1 \leq x_2 \leq 5 \\ & x_2 \in \mathbb{Z} \end{array}
$$

Examples of ”hidden” convex problems

$$
\begin{array}{llr}
\mathbf{P}: & \min & x_1 - 2x_2^2 \\
& \text{s.t.} & x_1^2 + x_2^2 \le 1 \\
& & x_2 = 0 \\
& & 0 \le x_1 \le 5
\end{array}
$$

Your first Optimization model

Variables	r: radius of the can's base
	h: height of the can
Objective	$2\pi rh + 2\pi r^2$ (minimize)
Constraints	$\pi r^2 h = V$
	$h > 0$
	$r > 0$

Relaxations

Relaxation of an Optimization problem

Consider an optimization problem

$$
\mathbf{P}: \qquad \begin{array}{rl} \min & f_0(x) \\ \text{s.t.} & f_1(x) \leq b_1 \\ & f_2(x) \leq b_2 \\ & \vdots \\ & f_m(x) \leq b_m, \end{array}
$$

Let us denote F the set of points x that satisfy all constraints:

$$
F = \{x \in \mathbb{R}^n : \begin{array}{l} f_1(x) \leq b_1, \\ f_2(x) \leq b_2, \\ \vdots \\ f_m(x) \leq b_m\} \end{array}
$$

So we can denote $\mathbf{P}: \min\{f_0(x) : x \in F\}$ for short.

Relaxation of an Optimization problem

Consider a problem $\mathbf{P} : \min\{f_0(x) : x \in F\}$.

A problem $\mathbf{P}' : \min\{f_0'(x) : x \in F'\}$ is a **relaxation** of $\mathbf{P}$ if:

- $F' \supseteq F$
- $f_0'(x) \leq f_0(x)$ for all $x \in F$.[3]

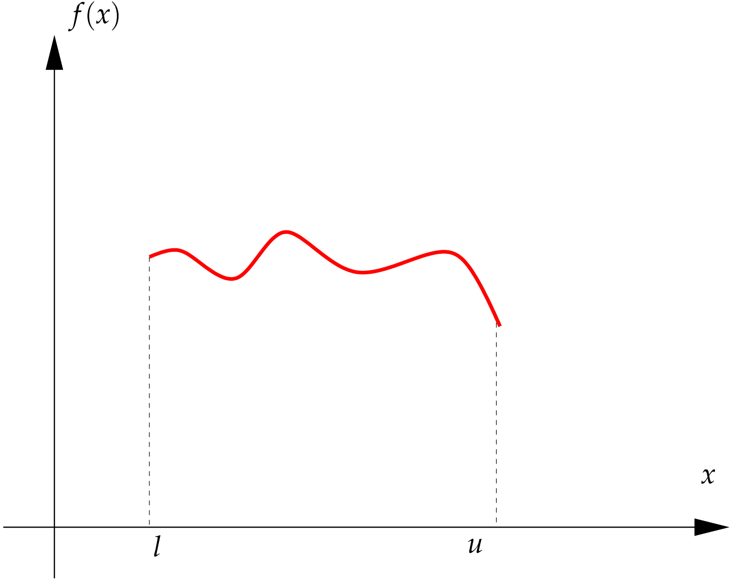
If $\mathbf{P}'$ is a relaxation of a problem $\mathbf{P}$, then the global optimum of $\mathbf{P}'$ is $\leq$ the global optimum of $\mathbf{P}$.

[3]We don't care what $f_0'(x)$ is outside of F.

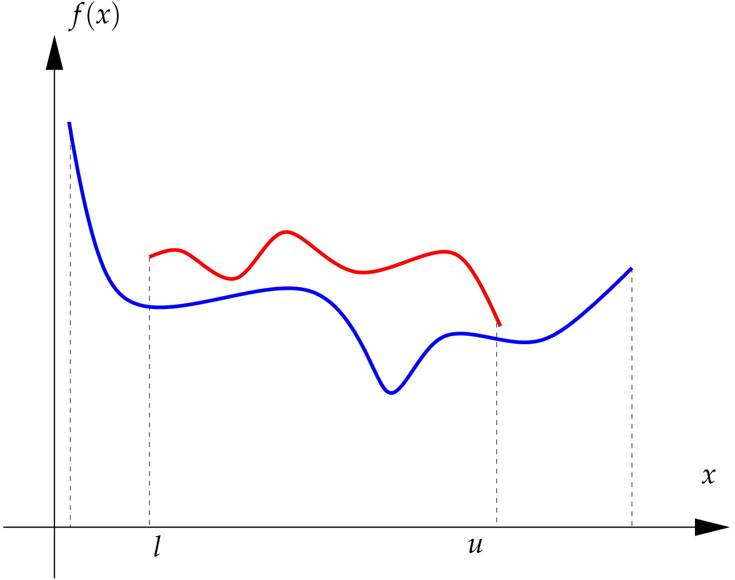
Examples

- $\min\{f(x) : -1 \leq x \leq 1\}$ is a relaxation of $\min\{f(x) : x = 0\}$
- $\min\{f(x) : -1 \leq x \leq 1\}$ is a r. of $\min\{f(x) : 0 \leq x \leq 1\}$
- $\min\{f(x) : -1 \leq x \leq 1\}$ is not a r. of $\min\{f(x) : -2 \leq x \leq 1\}$
- $\min\{f(x) : g(x) \leq b\}$ is a r. of $\min\{f(x) : g(x) \leq b - 1\}$
- $\min\{f(x) - 1 : g(x) \leq b\}$ is a r. of $\min\{f(x) : g(x) \leq b\}$

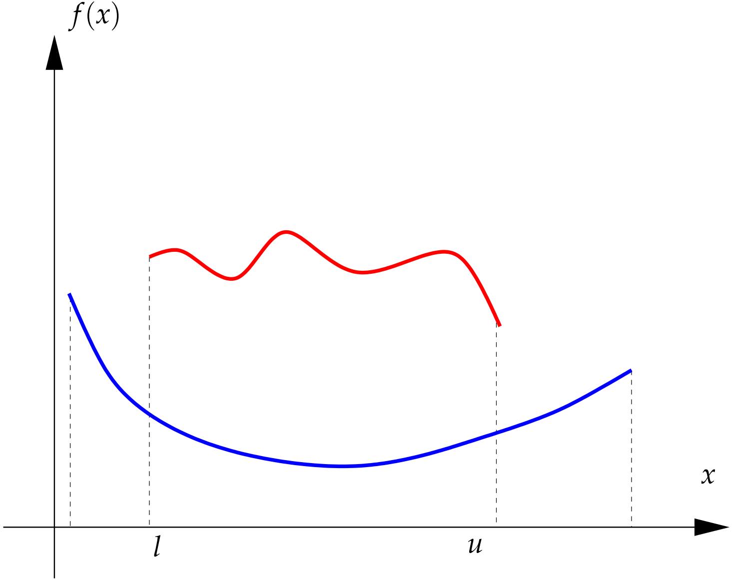
Relaxation of an Optimization problem



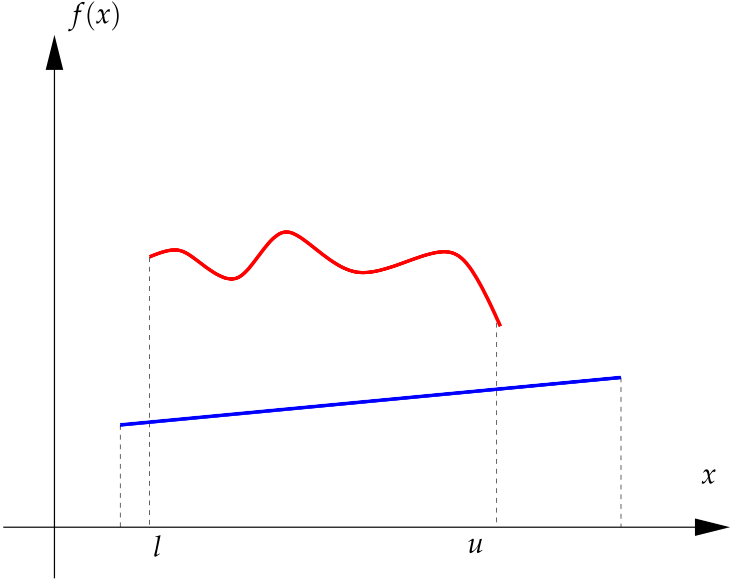
Relaxation of an Optimization problem



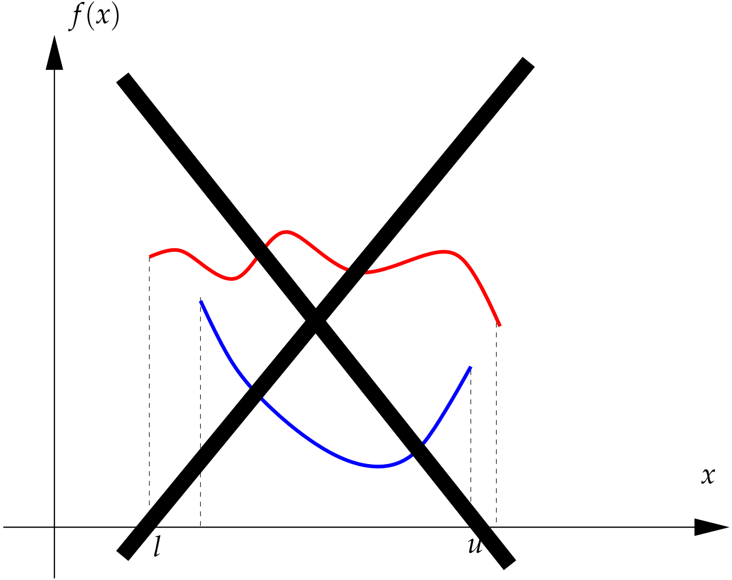
Relaxation of an Optimization problem



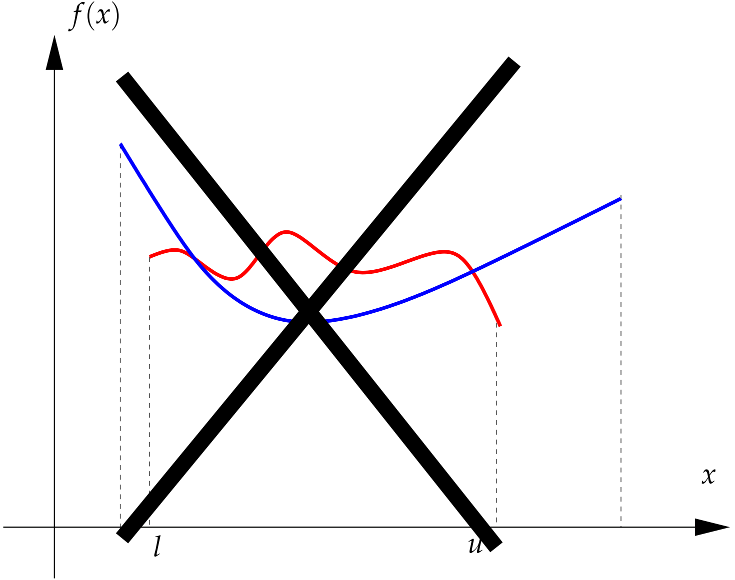
Relaxation of an Optimization problem



Relaxation of an Optimization problem



Relaxation of an Optimization problem



Relaxation of an Optimization problem

Consider again a problem
$\mathbf{P}$: $\min\{f_0(x) : f_1(x) \le b_1, f_2(x) \le b_1, \ldots, f_m(x) \le b_m\}$, or
$\mathbf{P}$: $\min\{f_0(x) : x \in F\}$ for short.

- deleting a constraint from $\mathbf{P}$ provides a relaxation of $\mathbf{P}$.
- adding a constraint $f_{m+1}(x) \le b_{m+1}$ to a problem $\mathbf{P}$ does the opposite:

$$F'' = \{x \in \mathbb{R}^n : \begin{array}{l} f_1(x) \le b_1, \\ f_2(x) \le b_2, \\ \ldots, \\ f_m(x) \le b_m, \\ f_{m+1}(x) \le b_{m+1}\} \subseteq F \end{array}$$

and therefore

$$\min\{f_0(x) : x \in F''\} \quad \ge \quad \min\{f_0(x) : x \in F\}$$

$\boxed{\textbf{Upper \& Lower bounds}}$

Lower and upper bounds

Consider an optimization problem $\mathbf{P} : \min\{f_0(x) : x \in F\}$:

- for any feasible solution $x \in F$, the corresponding objective function value $f_0(x)$ is an upper bound.
- the most interesting upper bounds are the local optima.
- a lower bound of $\mathbf{P}$ is instead a value z such that

$$z \leq \min\{f_0(x) : x \in F\}.$$

Upper vs. Lower bounds

Situation #1:

You: "We found a solution that will only cost 372,000 $."

Boss: "Ok, that sounds good."

Upper vs. Lower bounds

Situation #1:

You: “We found a solution that will only cost 372,000 \$.”

Boss: “Ok, that sounds good.”

Situation #2:

You: “We found a solution that will only cost 372,000 \$.”

Boss: “That’s too much, find something better.”

Upper vs. Lower bounds

Situation #1:

You: "We found a solution that will only cost 372,000 $."

Boss: "Ok, that sounds good."

Situation #2:

You: "We found a solution that will only cost 372,000 $."

Boss: "That's too much, find something better."

...

You: "We found another solution that costs 354,000 $."

Boss: "Can't you do better than that?"

Upper vs. Lower bounds

Situation #1:

You: "We found a solution that will only cost 372,000 $."

Boss: "Ok, that sounds good."

Situation #2:

You: "We found a solution that will only cost 372,000 $."

Boss: "That's too much, find something better."

...

You: "We found another solution that costs 354,000 $."

Boss: "Can't you do better than that?"

You: "I can try again, but here's the proof that we can't go below 351,500."

Boss: "Ok then, that's a good solution."

What relaxations are for

- If $\mathbf{P}'$ is a relaxation of a problem $\mathbf{P}$, then the global optimum of $\mathbf{P}'$ is $\leq$ the global optimum of $\mathbf{P}$.
- Hence, any relaxation $\mathbf{P}'$ of $\mathbf{P}$ provides a lower bound on $\mathbf{P}$.

$\Rightarrow$ If a problem $\mathbf{P}$ is difficult but a relaxation $\mathbf{P}'$ of $\mathbf{P}$ is easier to solve than $\mathbf{P}$ itself, we can still try and solve $\mathbf{P}'$: (i) we get a lower bound and (ii) the solution of $\mathbf{P}'$ may help solve $\mathbf{P}$.

To recap: the Knapsack problem

At a flea market in Rome, you spot n objects (old pictures, a vessel, rusty medals…) that you could re-sell in your antique shop for about double the price.

- You want these objects to pay for your flight ticket to Rome, which cost C.
- Also, you don't want a heavy backpack, so you want to buy the objects that will load it as little as possible.

How do you solve this problem?

The Knapsack problem

Each object $i = 1, 2, \dots, n$ has a price p_i and a weight w_i.

- Variables: one variable x_i for each $i = 1, 2 \dots, n$. This is a "yes/no" variable: either you take the i-th object or not.

The Knapsack problem

Each object $i = 1, 2, \ldots, n$ has a price p_i and a weight w_i.

- Variables: one variable x_i for each $i = 1, 2 \ldots, n$. This is a "yes/no" variable: either you take the i-th object or not.
- Constraint: total revenue must be at least C

The Knapsack problem

Each object $i = 1, 2, \ldots, n$ has a price p_i and a weight w_i.

- Variables: one variable x_i for each $i = 1, 2 \ldots, n$. This is a "yes/no" variable: either you take the i-th object or not.
- Constraint: total revenue must be at least C

  (As you'll double the price when selling them at your store, the revenue for each object is exactly p_i)

The Knapsack problem

Each object $i = 1, 2, \ldots, n$ has a price p_i and a weight w_i.

- Variables: one variable x_i for each $i = 1, 2 \ldots, n$. This is a "yes/no" variable: either you take the i-th object or not.
- Constraint: total revenue must be at least C

  (As you'll double the price when selling them at your store, the revenue for each object is exactly p_i)
- Objective function: the total weight

Your first (non-trivial) optimization model

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & \sum_{i=1}^{n} p_i x_i \geq C \\ & x_i \in \{0,1\} \quad \forall i = 1, 2, \ldots, n \end{array}$$

Is it convex?

Relaxing the Knapsack problem

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & x_i \in \{0,1\} \quad \forall i = 1, 2, \ldots, n \end{array}$$

Relaxing the Knapsack problem

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & x_i \in \{0,1\} \quad \forall i = 1, 2, \ldots, n \end{array}$$

This relaxation would give us $x_i = 0$ for all $i = 1, 2, \ldots, n$, and a lower bound of $\sum_{i=1}^{n} w_i x_i = 0$. Not so great…

Relaxing the Knapsack problem

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & x_i \in \{0,1\} \quad \forall i = 1,2,\ldots,n \end{array}$$

This relaxation would give us $x_i = 0$ for all $i = 1,2,\ldots,n$, and a lower bound of $\sum_{i=1}^{n} w_i x_i = 0$. Not so great…

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & \sum_{i=1}^{n} p_i x_i \geq C \\ & 0 \leq x_i \leq 1 \quad \forall i = 1,2,\ldots,n \end{array}$$

Relaxing the Knapsack problem

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & x_i \in \{0,1\} \quad \forall i = 1, 2, \ldots, n \end{array}$$

This relaxation would give us $x_i = 0$ for all $i = 1, 2, \ldots, n$, and a lower bound of $\sum_{i=1}^{n} w_i x_i = 0$. Not so great. . .

$$\begin{array}{ll} \min & \sum_{i=1}^{n} w_i x_i \\ & \sum_{i=1}^{n} p_i x_i \geq C \\ & 0 \leq x_i \leq 1 \quad \forall i = 1, 2, \ldots, n \end{array}$$

- Relaxing integrality of the variables gives a relaxation where we admit fractions of objects.

$\approx$ we pulverized all objects and took some spoonfuls of each

- It doesn't make sense, but it's a relaxation, and it **does** give us a better lower bound.