ECON 508-A

Finite Dimensional Optimization

Juan Ignacio Vizcaino

Washington University in St. Louis

July 29 - August 16 2019

Disclaimer.

These slides are based on the set of notes on Finite Dimensional Optimization by John Nachbar.

Section 1: Introduction

General form of an optimization problem:

$$\max_{\{x\}} \quad f(x)$$

$$\text{subject to}$$

$$x \in X$$

$x = (x_1, x_2, \ldots, x_N)$: N-dimensional vector of choice variables.

$f(x) = f(x_1, \ldots, x_N)$: real valued function called objective function.

$X \subseteq R^N$: constraint set, feasible set or opportunity set.

An optimization problem in $\mathbb{R}^N$ is one where, given function $f : \mathbb{R}^N \to \mathbb{R}$, we choose the value of x over a given set $X \subseteq \mathbf{R}^N$ that maximizes or minimizes the value of f.

Alternative notation:

$$\max_{\{x \in X\}} \; f(x) \qquad \qquad \min_{\{x \in X\}} f(x)$$

or

$$\max \{f(x) \mid x \in X\} \qquad \qquad \min \{f(x) \mid x \in X\}$$

Given f and X as above,

The set of attainable values of f on X, or **image** of X **under** f, denoted $f(X)$, is defined by

$$f(X) = \{\omega \in \mathbb{R} \mid \exists \text{ some } x \in X \text{ such that } f(x) = \omega\}$$

The **interior** of X is the set defined by

$$int(X) = \{x \in X \mid \exists \text{ some } \varepsilon > 0 \text{ such that } N_\varepsilon(x) \subseteq X\}$$

A solution to the problem $\max\{f(x) \mid x \in X\}$ or **maximizer** of f on X is a point $x^\star$ s.t.

$$f(x^\star) \geq f(x) \; \forall \, x \in X$$

A solution to the problem $\min\{f(x) \mid x \in X\}$ or **minimizer** of f on X is a point $x^\star$ s.t.

$$f(x^\star) \leq f(x) \; \forall \, x \in X$$

Local vs Global Solutions

$x^\star \in X$ is a **global maximum** of f in X iff $f(x^\star) \geq f(x)\ \forall\, x \in X$

$x^\star \in X$ is a **local maximum** of f in X iff there is an $\varepsilon > 0$ such that $N_\varepsilon(x^\star) \subseteq X$ and $f(x^\star) \geq f(x)$ for any $x \in N_\varepsilon(x^\star) \cap X$

$x^\star \in X$ is an **interior local maximum** of f in X iff there is an $\varepsilon > 0$ such that $N_\varepsilon(x^\star) \subseteq X$ and $f(x^\star) \geq f(x)\ \forall\, x \in N_\varepsilon(x^\star)$

The defnitions for minima are analogous, but with inequalities reversed.

Set of Solutions

The **set of solutions** to a max problem is denoted

$$\text{argmax}\Big(f(x) \mid x \in X\Big) = \big\{x \in X \mid f(x) \geq f\left(x'\right) \forall\, x' \in X\big\}$$

$\text{argmax}\Big(f(x) \mid x \in X\Big)$ can have more than one element.

Example 1
Let $X = [-1, 1]$, and $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$.
Maximizing f on C has two solutions, $x = -1$ and $x = 1$.
$\text{argmax}\Big(f(x) \mid x \in X\Big) = \{-1, 1\}$.

Example 2
Let $X = \mathbb{R}_+$, and $f : \mathbb{R}_+ \to \mathbb{R}$ be $f(x) = x^2$.
The problem of maximizing f on X has no solution.
$\text{argmax}\Big(f(x) \mid x \in X\Big) = \{\varnothing\}$

Existence of a Solution

A sufficient condition for existence of a solution is given by the **Weierstrass Theorem**.

Theorem 1.1 Weierstrass Theorem

If f is continuous and X is closed and bounded (hence compact) and nonempty, then there exist a global maximum and a global minimum.

Weierstrass is sufficient but not necessary for existence of a solution.

Example

$X = (0, 2]$ and $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$.

Consumer's Utility Maximization Problem

$$\max u(x)$$
$$\text{s.t.}$$
$$p \cdot x \leq m$$
$$x \geq 0$$
$$p \in \mathbb{R}^{l}_{+1}$$
$$x \in \mathbb{R}^{l}$$

Does a solution exist?

The constraint set $X = B(p, m) = \left\{x \in \mathbb{R}^{l} : p \cdot x \leq m\right\}$ is compact. So if $u(\cdot)$ is continuous, then a solution exists (by Weierstras's Theorem).

Firm's Cost Minimization Problem

$$\begin{aligned} \min{}_{\text{s.t.}}\ & f(x) \geq y \\ \text{s.t. } & f(x) \geq y \\ & x \geq 0 \\ & \mathbb{R} \in \mathbb{R}^l_{++}, y \in \mathbb{R}_{++} \\ & x \in \mathbb{R}^l \end{aligned}$$

Does a solution exist?

The constraint set *is not* compact.
If the constraint set is not empty, $\exists\, x^\star$ s.t. $f(x^\star) \geq y$.
Compactify the constraint set: $\{x : p \cdot x \leq p \cdot x^* \text{ and } f(x) \geq y\}$.
Since the objective function is continuous, there exists a solution.

Theorem 1.2

Let $-f$ denote the function whose value at any x is $-f(x)$. Then x is a maximum of f on X iff it is a minimum of $-f(x)$ on X; and z is a minimum of f on X iff z is a maximum of $-f(x)$ on X.

Theorem 1.3

Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be a strictly increasing function, that is, a function such that

$$y > y' \Rightarrow \varphi(y) > \varphi\left(y'\right)$$

Then x is a maximum of f on X iff x is a maximum of the composition of φ and f on X; and z is a minimum of f on X iff z is a minimum of $\varphi \circ f$ on X.

REMARK: it suffices that φ be a strictly increasing function on the set $f(X)$: That is, that φ only satisfies $\varphi(y) > \varphi(y')$ for all $y > y'$

Section 2: Finite Dimensional Constrained Optimization

Roadmap

We are going to split Maximization Theory into three parts:

Part I Provide necessary conditions: *conditions that solutions to maximization problems must satisfy.*

Part II Provide sufficient conditions: *conditions that guarantee that a candidate solution is indeed a maximum (or a minimum, or saddle point).*

Part III Convex Programming: *extension to situations where the functions are not differentiable.*

We are going to focus on maximization. Minimization is equivalent to maximizing $-f$. The results for maximization have easy corollaries for minimization.

Section 2.1: Finite Dimensional Optimization - Necessary Conditions

Let $C \subseteq X \subseteq \mathbb{R}^N$, $f : X \to \mathbb{R}$, and $x^* \in C$.

Consider the Maximization Problem

$$\max_{\{x \in C\}} \; f(x)$$

The results here assume C is written in *standard form*.

Definition
A maximization problem is in *standard form* iff there are K functions $g_k : X \to \mathbb{R}$ such that C is the set of x such that $g_k(x) \leq 0$ for all k. For minimization problems $g_k(x) \geq 0$ for all k.

Mnemonic: Impose *ceilings* (*floors*) in maximization (minimization) problems, blocking further increases (decreases) in f via x.

Example

Expressing a Problem (P) in standard form is no big deal.

The Consumer's Utility Maximization Problem is often written as:

$$\max_{\{x\}} u(x)$$
$$\textit{subject to}:$$
$$p \cdot x \leq m,$$
$$x \geq 0.$$

In standard form, the problem is:

$$\max_{\{x\}} u(x)$$
$$\textit{subject to}:$$
$$p \cdot x - m \leq 0,$$
$$-x_1 \leq 0,$$
$$\vdots$$
$$-x_N \leq 0.$$

- Let x^* be feasible, so that $g_k(x^*) \leq 0$
- An inequality constraint is **binding** at a certain point x^* if $g_k(x^*) = 0$. If $g_k(x^*) < 0$ at x^* it is **slack**.
- Define J to be the set of **indices** of the binding constraints:

$$J = \{k \in \{1, \ldots, K\} : g_k(x^*) = 0\}$$

KKT states that if x^* is a local maximum, $J \neq \emptyset$, then, subject to a technical condition called *Constraint Qualification*, for each $k \in J$ there is a number $\lambda_k^* \geq 0$ such that,

$$\nabla f(x^*) = \sum_{k \in J} \lambda_k \nabla g_k(x^*)$$

We will call this the **KKT condition**.

- Q. What if $J = \emptyset$?

Assume $J \neq \emptyset$, and let A be the set that is positively spanned by the gradients of the binding constraints:

$$A = \left\{ x \in \mathbb{R}^N : \exists\, \lambda_k \geq 0 \ \text{ s.t. } x = \sum_{k \in J} \lambda_k \nabla g_k\left(x^*\right) \right\}$$

Notice that A is a cone. Why?

KKT thus says that, if $J \neq \emptyset$, then $\nabla f\left(x^*\right)$ **lies in the cone A**.

λ_k^* measures how much the objective function would be increased if constraint k were relaxed slightly.

KKT Geometric Intuition

Remember: If the gradient of f at point P is not the zero vector, it has the direction of greatest increase of f at P.

The intuition behind KKT is that if x^* is a local maximum but $\nabla f(x^*) \neq 0$ then moving in a direction that increases f must violate one of the binding constraints, which means moving in a direction that increases at least one of the binding g_k functions.

MF Constraint Qualification

We need some additional condition. Why?

Example

Let the domain be $\mathbb{R}$. Let $f(x) = x$ and let $g(x) = x^3$. Then the constraint set is $C = (-\infty, 0]$ and the solution $x^* = 0$.
KKT fails! $\nabla f(x^*) = 1$ while $\nabla g(x^*) = 0$, there is no $\lambda \geq 0$ such that $\nabla f(x^*) = \lambda \nabla g(x^*)$

Definition Mangasarian-Fromovitz Constraint Qualification (MF) holds at x^* if there exists a $v \in \mathbb{R}^N$ such that for all $k \in J$, $\nabla g_k(x^*) \cdot v < 0$.

MF is the most useful form of constraint qualification, but not the only one and not the weakest.

Theorem 2.1.1 Karush-Kuhn-Tucker

Let x^* be a local solution to a differentiable MAX in standard form. If $J = \emptyset$, then $\nabla f(x^*) = 0$. If $J \neq \emptyset$ and MF holds at x^*, then for every $k \in J$ there is a $\lambda_k^* \geq 0$ such that

$$\nabla f(x^*) = \sum_{k \in J} \lambda_k^* \nabla g_k(x^*).$$

This equality is called the KKT condition and the λ_k are called the KKT multipliers.

Proof

Claim: If x^* is feasible but the KKT condition does not hold then there is a $v \in \mathbb{R}^N$ such that $\nabla f(x^*) \cdot v > 0$ and, for any $k \in J$, $\nabla g_k(x^*) \cdot v < 0$.

- Since f is differentiable, $\nabla f(x^*) \cdot v$ equals the directional derivative of f at x^* in the direction v. Since $\nabla f(x^*) \cdot v > 0$, this implies that there is a $\gamma_f > 0$ such that for all $\gamma \in (0, \gamma_f)$, $f(x^* + \gamma v) - f(x^*) > 0$, hence $f(x^* + \gamma v) > f(x^*)$.
- By a similar argument, since $\nabla g_k(x^*) \cdot v < 0$ for any $k \in J$, there is a $\gamma_k > 0$ such that for all $\gamma \in (0, \gamma_k)$, $g_k(x^* + \gamma v) - g_k(x^*) < 0$, hence $g_k(x^* + \gamma v) < 0$.
- Finally, for any $k \notin J$, since $g_k(x^*) < 0$ and g_k is continuous (since it is differentiable), there is $\gamma_k > 0$ such that for all $\gamma \in (0, \gamma_k)$, $g_k(x^* + \gamma v) < 0$.

Let $\overline{\gamma}$ be the minimum of γ_f and the γ_k note that $\overline{\gamma} > 0$ since it is a minimum of a finite set of strictly positive numbers. Then for all $\gamma \in (0, \overline{\gamma})$

1. $f(x^* + \gamma v) > f(x^*)$.
2. $g_k(x^* + \gamma v) < 0$ for all k (if there are any constraints at all).

Thus, for all $\gamma \in (0, \overline{\gamma}), x^* + \gamma v$ is feasible and yields a higher value for the objective function. Therefore x^* is not a local solution.

It remains to show that the above claim is true.

Consider the case where $J \neq \emptyset$. The KKT condition is equivalent to requiring that

$$\nabla f(x^*) \in A = \left\{ x \in \mathbb{R}^N : \exists \lambda_k \geq 0 \text{ s.t. } x = \sum_{k \in J} \lambda_k \nabla g_k(x^*) \right\}$$

where A is the cone positively spanned by the gradients of the binding constraints.
By contraposition: if $\nabla f(x^*) \notin A$ then there exists a $v \in \mathbb{R}^N$ such that $\nabla f(x^*) \cdot v > 0$ and for all $k \in J, \nabla g_k(x^*) \cdot v < 0$.

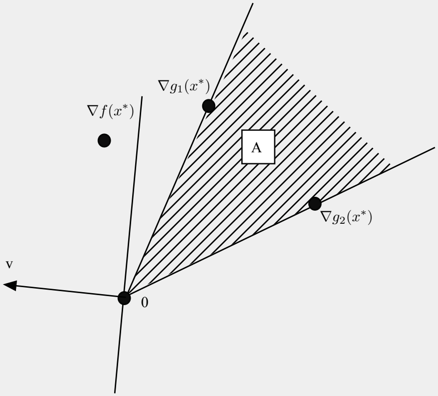
KKT Proof continued III

The cone A is closed and convex. By the Separating Hyperplane Theorem for Cones in the notes on Cones, there is a $\hat{v} \in \mathbb{R}^N$ such that $\nabla f(x^*) \cdot \hat{v} > 0$ and for all $a \in A, a \cdot \hat{v} \leq 0$.

By MF, there is a $\tilde{v} \in \mathbb{R}^N$ such that $\nabla g_k(x^*) \cdot \tilde{v} < 0$ for all $k \in J$. Take $\theta \in (0, 1)$ and let $v = \theta\tilde{v} + (1 - \theta)\hat{v}$.

- $\nabla f(x^*) \cdot v = \theta \nabla f(x^*) \cdot \tilde{v} + (1 - \theta)\nabla f(x^*) \cdot \hat{v}$, which is strictly positive for θ small even if $\nabla f(x^*) \cdot \tilde{v}$ is negative.
- $\nabla g_k(x^*) \cdot v = \theta \nabla g_k(x^*) \cdot \tilde{v} + (1 - \theta)\nabla g_k(x^*) \cdot \hat{v}$, which is strictly negative for any $\theta \in (0, 1)$.

KKT Proof: Geometric Intuition



Definition

A Maximization problem in standard form satisfies the Slater Condition iff there is a point x such that $g_k(x) < 0$ for all k.

- The Slater Condition is equivalent to requiring that the constraint set C has a non-empty interior.
- MF implies Slater. MF: there are points arbitrarily near x^* that are interior to the constraint set.

Theorem 2.1.2

In a differentiable MAX problem, at any feasible x^*, Slater is equivalent to MF if either

1. Each binding constraint function g_k is convex,
2. Each binding constraint function g_k is quasi-convex and $\nabla g_k(x^*) \neq 0$.

The Slater Condition: Example

- In economic applications, all constraints are often convex in MAX problems (or concave in MIN problems).
- Checking MF boils down to checking Slater, which is often trivial.

Example
Let the domain be $\mathbb{R}, f(x) = x, g(x) = x^3$.

Other Forms of Constraint Qualifications

Let S be the set of the gradients of the binding constraints, $S = \{\nabla g_k(x^*)\}_{k \in J}$.

Definition
Linearly independent constraint qualification (LI) holds at x^* iff S is linearly independent: if $\sum_{k \in J} \lambda_k \nabla g_k(x^*) = 0$ then $\lambda_k = 0$ for all $k \in J$. LI holds vacuously if $J = \emptyset$.

Definition
Positive linearly independent constraint qualification (PI) holds at x^* iff S is positive linearly independent: if $\sum_{k \in J} \lambda_k \nabla g_k(x^*) = 0$ and $\lambda_k \geq 0$ for all $k \in J$ then $\lambda_k = 0$ for all $k \in J$. PI holds vacuously if $J = \emptyset$.

Theorem 2.1.3

MF, PI, PC are equivalent.

Section 2.2: Finite Dimensional Optimization - Sufficient Conditions

- KKT states that, for a feasible point x^* and a set of binding constraints J there exists a $\lambda_k \geq 0$ for all $k \in J$ such that

$$\nabla f(x^*) = \sum_{k \in J} \lambda_k \nabla g_k(x^*).$$

- If no constraints are binding ($J = \emptyset$) then the KKT condition reduces to $\nabla f(x^*) = 0$.

- The KKT condition is **necessary**, but not always **sufficient**.

Example.
Let the domain be $\mathbb{R}$. Consider the functions $f_1(x) = -x^4$, $f_2(x) = x^4$, $f_3(x) = x^3$.

Definition

A function f is a *differentiably strictly increasing transformation* of a function $\hat{f}$ iff both functions have the same domain and there is a differentiable function h, with domain containing the image of $\hat{f}$, such that:

1. $f = h \circ \hat{f}$
2. $Dh(\hat{f}(x)) > 0$

- Any concave function is, trivially, a differentiably strictly increasing transformation of a concave function ($f = \hat{f}$ and h to be the identity, $h(y) = y$).

Theorem 2.2.1

Consider a differentiable MAX problem in standard form with objective function f.

1. If x^* is feasible, and the KKT condition holds at x^*, then x^* is a solution to the MAX problem if either,
   a f is concave (or is a differentiably strictly increasing transformation of a concave function), or
   b f is quasi-concave, $\nabla f(x^*) \neq 0$, and any binding constraint functions are quasi-convex.
2. If f is strictly quasi-concave, and if the constraint functions are quasi-convex, then the solution to the MAX problem is unique.

1. The proof is by contraposition. Suppose that there is a feasible x such that $f(x) > f(x^*)$. Let $v = x - x^*$. I claim that $\nabla f(x^*) \cdot v > 0$.
   (a) Suppose that f is concave. Then $f(x) \leq \nabla f(x^*) \cdot (x - x^*) + f(x^*)$, hence $0 < f(x) - f(x^*) \leq \nabla f(x^*) \cdot v$.

   Suppose that f is a differentiably strictly increasing transformation of $\hat{f}$, with $\hat{f}$ concave. Since $f(x) > f(x^*)$ iff $\hat{f}(x) > \hat{f}(x^*)$, the above argument implies $\nabla \hat{f}(x^*) \cdot v > 0$. Since, by the Chain Rule, $\nabla f(x^*) = Dh\left(\hat{f}(x^*)\right) \nabla f(x^*)$, it follows that $\nabla f(x^*) \cdot v = Dh\left(\hat{f}(x^*)\right) \nabla \hat{f}(x^*) \cdot v > 0$.

(b) Suppose that f is merely quasi-concave. Since f is continuous, there is an $\varepsilon > 0$ such that for any w on the unit sphere in $\mathbb{R}^N$, $f(x+\varepsilon w) > f(x^*)$. For any w on the unit sphere, and for any $\theta \in (0,1)$, quasi-concavity then implies that $f(x^*) \leq f(\theta(x+\varepsilon w)+(1-\theta)x^*) = f(x^* + \theta(x+\varepsilon w - x^*))$ or

$$f(x^* + \theta(x+\varepsilon w - x^*)) - f(x^*) \geq 0$$

Dividing by $\theta > 0$ and taking the limit as $\theta \downarrow 0$ implies that the directional derivative of f at x^* in the direction $x+\varepsilon w - x^*$ is non-negative. Since f is differentiable, this implies that,

$$\nabla f(x^*) \cdot (x+\varepsilon w - x^*) \geq 0$$

hence,

$$\nabla f(x^*) \cdot v + \varepsilon \nabla f(x^*) \cdot w \geq 0$$

. This holds for all w on the unit sphere. Since $\nabla f(x^*) \neq 0$, there is a w such that $\nabla f(x^*) \cdot w < 0$. The claim follows.

Theorem 2.2.1: Proof III

If $J = \emptyset$, then $\nabla f(x^*) \cdot v > 0$ implies that the KKT condition (which, in this case, is $\nabla f(x^*) = 0$) does not hold, and the proof follows by contraposition.
If $J \neq \emptyset$ then, by the KKT condition,

$$0 < \nabla f(x^*) \cdot v = \sum_{k \in J} \lambda_k \nabla g_k(x^*) \cdot v$$

which implies that there is at at least one $k \in J$ such that $\nabla g_k(x^*) \cdot v > 0$. But since g_k is differentiable, this implies that the directional derivative of g_k at x^* in the direction v is strictly positive. This then implies that for all $\theta \in (0, 1)$ sufficiently small, $g_k(x^* + \theta v) > g_k(x^*)$, hence $g_k(\theta x + (1 - \theta)x^*) > g_k(x^*)$. But since x is feasible, $g_k(x) \leq 0$, and since $k \in J, g_k(x^*) = 0$. Together, these inequalities imply that g_k is not quasi-convex. Again, the proof follows by contraposition.

2. Suppose that there is a feasible x with $f(x) = f(x^*)$. Then, by the definition of strict quasi-concavity, if $x \neq x^*$ then for any $\theta \in (0, 1), f(\theta x + (1 - \theta)x^*) > f(x^*)$. Since $\theta x + (1 - \theta)x^*$ is feasible (the constraint set is convex if the g_k are quasi-convex), it follows by contraposition that the maximum is unique.

Checklist for Optimization Problems.

The Theorem above, together with KKT Theorem, give a checklist for optimization problems.

Theorem 2.2.2

Consider a differentiable MAX problem in standard form, with objective function f. Let x^* be feasible. If

1. f is either (a) concave, or (b) a differentiably strictly increasing transformation of a concave function, or (c) quasi-concave with $\nabla f(x^*) \neq 0$,
2. every binding constraint (if any) is either (a) convex or (b) quasi-convex with $\nabla g_k(x^*) \neq 0$
3. the Slater condition holds,

then a necessary and sufficient for x^* to be a solution is that the KKT condition holds at x^*.

Consider the following problem,

$$\max_{x \in \mathbb{R}^N_+} \quad \prod_n x_n^{\alpha_n}$$
$$p \cdot x \leq m$$

with $\alpha_n \in (0, 1)$ for all n, $\sum_n \alpha_n = 1$, $p \in \mathbb{R}^N_{++}$, $m \in \mathbb{R}_{++}$.

1. The objective function is actually concave. Consider the following instead:

$$\hat{f}(x) = \sum_n \alpha_n \ln(x_n)$$

2. The constraints are convex (they are linear).
3. Slater holds: take $x_n = m/(2Np_n)$ for each n.

Section 2.3: Finite Dimensional Optimization - Convex Optimization

- For simplicity focus on MAX problems with a single variable $x \in \mathbb{R}$, and a single constraint $g : \mathbb{R} \to \mathbb{R}$
- Given a MAX problem, define $\mathcal{L} : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ by

$$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x)$$

- $\mathcal{L}$ is called the **Lagrangian**.
- (x^*, λ^*) is a **saddle point** of $\mathcal{L}$ iff

$$\mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x^*, \lambda)$$

  for all $x \in \mathbb{R}, \lambda \in \mathbb{R}_+$.
- (x^*, λ^*) is a saddle point iff (x^*, λ^*) maximizes $\mathcal{L}$ with respect to x and minimizes $\mathcal{L}$ with respect to to λ.
- We are going to study theorems that relate a saddle point to the KKT condition.

Theorem 2.3.1

Given a MAX problem in standard form and associated Lagrangian $\mathcal{L}$, if (x^*, λ^*) is a saddle point of $\mathcal{L}$ and $\mathcal{L}$ is differentiable then x^* is feasible and the **KKT** condition holds at x^*.

- Theorem 2.3.1 states that if (x^*, λ^*) is a saddle point, then the KKT condition holds at x^*.

Theorem 2.3.2

Given a MAX problem in standard form and associated $\mathcal{L}$, if (x^*, λ^*) is a saddle point of $\mathcal{L}$ then x^* solves MAX.

- If (x^*, λ^*) is a saddle point, then x^* is a maximum.

Saddle points and KKT: Theorems Continued.

- So far we know that if (x^*, λ^*) is a saddle point of $\mathcal{L}$, KKT holds and x^* solves the MAX problem.
- What conditions do we need ir order for (x^*, λ^*) to be a saddle point?

Theorem 2.3.3

Given a MAX problem in standard form and associated $\mathcal{L}$, suppose that x^* solves MAX. If f is concave, g is convex, and g satisfies the Slater condition, then there is an $\lambda^* \geq 0$ such that (x^*, λ^*) is a saddle point of $\mathcal{L}$.

Interpretation of Theorem 2.3.3.

- Interpreting Theorem 2.3.3 gives us intuition for solving problems in practice.
    1. If the constraint is convex and Slater holds then constraint qualification holds.
    2. By KKT Theorem, the KKT condition holds at x^*.
    3. If g is not binding ($g(x^*) < 0$), $\mathcal{L}(x^*, \lambda)$ which is strictly increasing in λ, is minimized at $\lambda^* = 0$.
    4. If g is binding ($g(x^*) = 0$), $L(x^*, \lambda) = f(x^*)$ and it is minimized at any $\lambda^* \geq 0$, and, in particular $\lambda^* = 0$.
    5. As for x, the KKT condition implies that $\nabla_x L(x^*, \lambda^*) = 0$. Since $\mathcal{L}$ is concave (since f is concave and g is convex), $\nabla_x L(x^*, \lambda^*) = 0$ is a sufficient condition for x^* to maximize $L(x, \lambda^*)$.
- In practice, we take First Order Conditions to find the x^* that satisfies KKT and *shut down* the role of λ by imposing combinations of $\lambda = 0$ and $g(x^*) \geq 0$) (interior solution) or $\lambda \geq 0$ and $g(x^*) = 0$ (corner solution), s.t. $\lambda \cdot g(x^*) = 0$.

Theorem 2.3.4

Given a differentiable MAX problem in standard form, if the objective function is concave, the constraint is convex, and Slater holds, then a necessary and sufficient condition for a feasible x^* to be a solution is that the **KKT** condition holds at x^*.

Proof.
If x^* is a solution, then Theorem 2.3.3 says that there is a $\lambda^* \geq 0$ such that (x^*, λ^*) is a saddle point of $\mathcal{L}$ and Theorem 2.3.1 then says that the KKT condition holds.
Conversely, if the KKT condition holds at x^* then the argument given above, following the statement of Theorem 2.3.3, λ^* such that (x^*, λ^*) is a saddle point of $\mathcal{L}$. Theorem 2 then says that x^* solves MAX.