# Some More Math for Economists: Concavity and Convexity

Spring 2023

# Convex Sets

Let $x^*$ be a solution to

$$\max f(x) \text{ s.t } g(x) \leq 0$$

Geometrically, the Lagrange multiplier conditions told us that

$$\{x : g(x) \leq 0\} \text{ and } \{x : f(x) \geq f(x^*)\}$$

are "tangent."

For nice convex sets, it looks like this tangency is sufficient.

# Convex Sets

**Definition (Convex Sets)**

A set $A \subseteq \mathbb{R}^n$ is convex if for any $\lambda \in [0, 1]$, any $x, y \in A$, $\lambda x + (1 - \lambda) y \in A$

# Hyperplanes

Fix a vector $d \in \mathbb{R}^n \setminus \{0\}$ and a constant $c \in \mathbb{R}$. These define a hyperplane:

$$H(d, c) = \{x : d \cdot x = c\}$$

where $d$ determines the orientation, and c determines the position.

# Supporting and Separating Hyperplanes

**Definition (Supporting Hyperplane)**

$H(d, c)$ is a supports set $A$ through point $x \in A$ if $d \cdot x = c$ and $d \cdot y \leq c$ for all $y \in A$

**Definition (Separating Hyperplane)**

$H(d, c)$ (strictly) separates sets $A$ and $B$ if for all $a \in A$, $b \in B$, $d \cdot a \leq (<) c \leq (<) d \cdot b$

# Separating Hyperplane Theorem

**Theorem (Separating Hyperplane Theorem)**

*Let $A, B$ be non-empty, closed, disjoint, convex sets. Suppose $A$ is compact. Then there exists a hyperplane $H(d, c)$ s.t. $H(d, c)$ strictly separates $A$ and $B$.*

# Separation and Optimization

Suppose the feasible set:

$$F = \{x : g(x) \leq 0\}$$

is convex, closed, and compact. Suppose that for every $k$

$$UC(k) = \{x : f(x) \geq k\}$$

are convex and closed.

We can then provide a simple geometric characterization of optimality.

# Separation and Optimization

Rewording this observation:

**Theorem**

*Consider a strictly increasing function $f : \mathbb{R}^M \to \mathbb{R}$ with closed, convex $UC(k)$ and a set $F \subseteq \mathbb{R}^M$ closed, compact and convex, $x^*$ solves*

$$\max_{x \in F} f(x)$$

*if and only if there exists some $d \in \mathbb{R}^n \setminus \{0\}$ such that both of the following are true:*

- *$x^*$ solves $\max d \cdot x$ s.t. $x \in F$.*
- *$x^*$ solves $\min d \cdot x$ s.t $x \in UC(f(x^*))$.*

This gives us a first necessary and sufficient condition for optimality.

# Proof

(If:) Let $x^*$ be a maximum.

- $x^*$ cannot be an interior point of $F$ or $UC(f(x^*))$.
- Therefore, there exists a hyperplane $H(d, d \cdot x^*)$ that separates $F$ and $UC(f(x^*))$.
- So $x^*$ solves both the maximization and minimization problem

# Proof

(If:) Let $x^*$ be a maximum.

- $x^*$ cannot be an interior point of $F$ or $UC(f(x^*))$.
- Therefore, there exists a hyperplane $H(d, d \cdot x^*)$ that separates $F$ and $UC(f(x^*))$.
- So $x^*$ solves both the maximization and minimization problem

(Only if) Suppose that for some feasible $x^*$ there exists a $d$ s.t. $x^*$ solves both the maximization and minimization problem.

- Then the hyperplane $H(d, d \cdot x^*)$ separates $F$ and $UC(f(x^*))$.
- So any $x$ s.t. $f(x) > f(x^*)$ is not an element of $F$. So $x^*$ is a maximizer.

# Separation

Our primitives were functions, not sets.

**Definition (Upper and Lower Contour sets)**

The upper contour sets of $f : \mathbb{R}^n \to \mathbb{R}$ are the sets $UC(k) = \{x : f(x) \geq k\}$ for each $k$. The lower contour sets are $LC(k) = \{x : f(x) \leq k\}$.

**Definition (Quasiconvexity)**

A function $f : X \to \mathbb{R}$ is called quasiconvex (quasiconcave) if its lower (upper) contour sets are convex.

# Quasiconcavity

Some alternative formulations of Quasiconcavity:

**Theorem**

*$f : X \to \mathbb{R}$ with convex domain $X$ is quasiconcave if and only if for any $x, y \in X$, $\lambda \in (0,1)$, $f(\lambda x + (1-\lambda)y) \geq \min(f(x), f(y))$*

**Theorem**

*$f : X \to \mathbb{R}$, $X$ convex, $f$ differentiable is quasiconcave iff*

$$f(y) \geq f(x) \Rightarrow Df(x)(y - x) \geq 0$$

# Quasiconcavity

So now we know that if $g$ is quasiconvex and $f$ is quasiconcave and both are continuous then any $d$ where we can jointly solve the pair of optimization problems identify the maximizer.

- $x^*$ solves $\max d \cdot x$ s.t. $g(x) \leq 0$.
- $x^*$ solves $\min d \cdot x$ s.t $f(x) \geq f(x^*)$.

Let's connect these back to KKT conditions.

# Quasiconcavity

Suppose $f$ and $g$ are differentiable, and that $x^*$ is a satisfies the KKT conditions and $\nabla f(x^*) \neq 0$, $Dg(x^*)$ satisfies the rank condition. Guess $d = \nabla f(x^*)$, we can use this to show that $x^*$ is a maximizer.

- We can write the maximization problem as

  $$\max \lambda' Dg(x^*) \cdot x \text{ s.t. } g(x) \leq 0$$

  which by complementary slackness, $\lambda$ positive, and quasiconvexity of $g(x)$ is solved by $x^*$.
- Quasiconcavity of $f$ directly implies that $x^*$ solves the minimization problem.

# Quasiconcavity and KKT

We could similarly show that this pair of problems imply the KKT conditions

- If $d \neq \nabla f(x^*)$ then $x^*$ doesn't solve the minimization problem.
- If $d \neq \lambda' Dg(x^*)$ for some $\lambda$ that satisfies complementary slackness then $x^*$ does not solve the maximization problem.

# Linear Production

Consider the following simple model of a firm.

- The firm sells $x \in \mathbb{R}^n$ units of each output according to price vector $p$.
- The firm has access to a set of production processes described by matrix $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^m$, where $c$ is the amount of inputs it has.
- So the firm solves:

$$\max_{x \in \mathbb{R}^n_+} p \cdot x$$
$$\text{s.t. } Ax \leq c$$

# Linear Production

$$\max_{x \in \mathbb{R}^n_+} p \cdot x$$
$$\text{s.t. } Ax \leq c$$

This is a linear program.

- Each row of the matrix describes a different way the firm can use inputs to produce outputs.
- There's a massive literature on linear programming, I mostly leave this for future courses.

# Linear Production

The KKT conditions are especially simple here. Let $(A)_i$ denote the kth column of $A$. Accounting for non-negativity we get

$$\begin{aligned} x_i(p_i - \lambda'(A)_i) &= 0 \text{ for all } i \in \{1, 2, \ldots n\} \\ \lambda_k(Ax - c)_k &= 0 \text{ for all } k \in \{1, 2, \ldots m\} \\ \lambda, x &\geq 0 \end{aligned}$$

We often talk about interpreting $\lambda$ as a "shadow price", this can be made very clear in this problem.

# Linear Production

To see this, consider the following alternative program:

$$\min_{y \in \mathbb{R}^m_+} y \cdot c$$
$$\text{s.t. } y'A \geq p$$

We can interpret $y$ as the price to rent each technology.

# Linear Production

The solution to the dual problem puts a price on each production technology.

The dual's KKT conditions are

$$\begin{aligned} \mu_i(p_i - y'(A)_i) &= 0 \\ y_k(A\mu - c)_k &= 0 \\ y, \mu &\geq 0 \end{aligned}$$

These look familiar. The multipliers in the original problem are exactly these rental prices in the dual problem.

- The original problem can in some sense be reformulated as solving for a "price" for each constraint.
- In fact, note that

  $$\begin{aligned} p \cdot x &= \lambda' A \cdot x \\ &= y' A \cdot \mu \\ &= y \cdot c \end{aligned}$$

  at the optimum of the respective programs.

# Concavity

Quasiconcavity is a bit awkward. Let's define a stronger property

**Definition (Concavity)**

A function $f : X \to \mathbb{R}$, $X$ convex, is concave if for any $\lambda \in (0, 1)$, $f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$.

# Concavity

**Theorem**

*Suppose $f$ is twice continuously differentiable. The following are equivalent:*

- $f(x)$ *is concave.*
- $f(y) \leq f(x) + \nabla f(x) \cdot (y - x)$.
- $D^2 f(x)$ *is negative semidefinite.*

These give us a natural economic interpretation for concave functions.

# Concavity and Uniqueness

Suppose $f : X \to \mathbb{R}$, $X$ convex, $f$ strictly concave. Then $f$ has a unique maximum.

- Suppose not. Then there exist maximizers $x^*, y^*$ s.t. $f(x^*) = f(y^*)$ and $x^* \neq y^*$.
- But, by strict concavity $f(\lambda x^* + (1-\lambda)y^*) > \lambda f(x^*) + (1-\lambda) f(y^*) > f(x^*)$.

This property holds under strict quasiconcavity as well.

# Necessary and Sufficient Conditions

We can use concavity to get simple necessity condition for KKT conditions

**Theorem (Slater's Condition)**

*The KKT conditions are necessary if $f$ is concave, each constraint is convex and there exists an $x$ where $g(x) << 0$. We have strong duality,*

$$\min_{\lambda \geq 0} L(\lambda) = \max_{x \in X} f(x) \text{ s.t. } g(x) \leq 0.$$

and a sufficient condition

**Theorem**

*The KKT conditions are sufficient for a maximum if $\nabla f(x) \neq 0$ for all feasible $f$, $f$ is quasiconcave and the constraints are quasiconvex.*

# KKT-Sufficiency

**Theorem (KKT sufficiency)**

*Suppose $\nabla f(x) \neq 0$ for all feasible $x$, $f$ quasiconcave, $g_i$ quasiconvex for all $i \in \{1, 2, \ldots m\}$. Then any point satisfying the KKT conditions is a global max.*

Proof:

- Fix a $(x^*, \lambda^*)$ that satisfies the KKT conditions. For any $y$

  $$\nabla f(x^*) \cdot (y - x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) \cdot (y - x^*) = 0$$

  By quasiconvexity

  $$\lambda_i^* \nabla g_i(x^*) \cdot (y - x) \leq 0$$

  since for any feasible $y$, if $g_i$ binds at $x^*$, $g_i(y) \leq g_i(x^*)$ and $g$'s are quasiconvex.
- So $\nabla f(x^*)(y - x^*) \leq 0$, and thus $f(x^*) \geq f(y)$ by quasiconcavity.

# What does Concavity Capture

- Concavity is a natural assumption to place on utility or production.
- Convexity is a natural assumption to place on costs.
- Concavity is a natural assumption use to model preferences over risky alternatives

  It implies an decision maker with concave utility over outcomes of a lottery $U(x)$ will always take the sure thing over the lottery
  $$U(E(X)) \geq E(U(X))$$
- Many endogenous objects are naturally concave or convex. The value of information is convex, the expenditure function is convex

# Portfolio Choice

A standard simplification in finance is that decision maker's preferences depend only on the mean and variance of their returns. This leads to the natural problem of finding the minimum variance portfolio

Suppose there are two risky assets, $\mu = (\mu_1, \mu_2)$ are their means, $\Sigma$ is their variance, covariance matrix. Assume they are not perfectly correlated There's also a riskless asset (s) with return $\mu_3$. Let's solve

$$\min_{(x,s)\in\mathbb{R}^3} x'\Sigma x$$

$$\text{s.t. } \mu_3 s + \mu \cdot x \geq M$$

$$s + x_1 + x_2 \leq 1$$

where $\mu_3 < \mu_2 < \mu_1$ and $\sigma_2 < \sigma_1$. The interesting case is $M > \mu_3$.

Note that the objective is strictly convex and the constraints are linear, so we can use KKT conditions here.

# Portfolio Choice

The KKT conditions give us

$$
\begin{aligned}
2\sigma_1^2 x_1 + 2\sigma_{12} x_2 &= \lambda_1 \mu_1 - \lambda_2 \\
2\sigma_2^2 x_2 + 2\sigma_{12} x_1 &= \lambda_1 \mu_2 - \lambda_2 \\
\lambda_1 \mu_3 - \lambda_2 &= 0
\end{aligned}
$$

as well as complementary slackness and positivity of the multiplier.

# Portfolio Choice

Simplifying a bit:

$$\sigma_1^2 x_1 + \sigma_{12} x_2 = \lambda_1(\mu_1 - \mu_3)$$
$$\sigma_2^2 x_2 + \sigma_{12} x_1 = \lambda_1(\mu_2 - \mu_3)$$

$\lambda_1 = 0$ would only be possible if the assets were perfectly correlated, so both constraints bind and the FOCs tell us that

$$\frac{x_2}{x_1} = \frac{\sigma_1^2 + \rho\sigma_1\sigma_2 \frac{\mu_1 - \mu_3}{\mu_2 - \mu_3}}{\sigma_2^2 \frac{\mu_1 - \mu_3}{\mu_2 - \mu_3} + \rho\sigma_1\sigma_2}$$

where $\rho$ is the correlation.

# Portfolio Choice

Some things to note:

- $x_1$ must always be positive to satisfy our constraints. The sign of this fractions determines whether you buy or sell $x_2$.
- The proportion of $x_1$ relative $x_2$ is independent of the target mean.
- The optimal portfolio is a mixture of the safe asset and a risky portfolio whose composition doesn't change.
- The mean/variance frontier has a nice structure since the optimal composition of the risky portfolio is fixed.
  - Consider the portfolio given by solving

$$x_2 = \frac{\sigma_1^2 + \rho\sigma_1\sigma_2\frac{\mu_1-\mu_3}{\mu_2-\mu_3}}{\sigma_2^2\frac{\mu_1-\mu_3}{\mu_2-\mu_3} + \rho\sigma_1\sigma_2} x_1$$

$$x_1 + x_2 = 1$$

and let $\mu_y$ denote it's mean and $\sigma_y$ denote it's variance.

So we can reformulate the minimization problem as

$$\min_{y,s\in\mathbb{R}_+} \sigma_y^2 y^2 \text{ s.t. } \mu_3 s + \mu_y y = M \text{ and } s + y = 1$$

This problem is very easy.

Varying $M$ tells us the lowest variance portfolio we can get for each target mean. The set of feasible mean/variance combinations is thus

$$Variance \geq \left(\frac{M-\mu_3}{\mu_y-\mu_3}\right)^2 \sigma_y^2.$$

So the problem of maximizing a utility function that only depends on the mean and variance is a maximization problem with a convex feasible set.