Assignment #1 for **Managerial Economics**
Spring 2019

Due date: Mon. Feb. 18, 2019

Required Readings

- The optimization and comparative statics in the problems and notes below.
- Kreps, Chapters 11 and 12 on production and costs.
- Kreps, Chapters 18 and 19 on expected utility theory and diversification.
- Klein, Coffee, Partnoy, Chapter 4.I and 4.II.

Recommended Readings

- The Amir survey article on supermodularity. We will be using the simple versions of his results extensively in this problem set. The material we need and use is covered, in a rather detailed way, below.
- The Hayek article, sections III and IV, on the uses of knowledge and information in society. We will use these insights in the next problem set, when we study transfer pricing within a firm or organization.
- The Khan article on the history of the rights of married women to sign legally binding contracts. This is a wonderful example of changes in legal arrangements leading to changes in behavior.
- The next assignment will include optimization treatments of examples found in Klein et al. Ch. 1, so it may be worth looking at these ahead of time.

1. ON UTILITY FUNCTIONS AND 'PURPOSEFUL BEHAVIOR'

It is worth reading pages ix-xv in the Preface to Kreps's textbook, especially how he describes "purposeful behavior." The essential model that we, as economists, use to describe purposeful behavior is the idea of utility maximizing behavior. The idea is that, faced with a set of alternatives, $X = \{x_1, x_2, \ldots, x_n\}$, the person making the decision chooses the x in the set that has the highest value of $u(x)$, where $u : X \to \mathbb{R}$ is a function measuring the utility of the options. Notationally, x^* is the optimal choice(s), and it is the solution to the problem

$$(1) \qquad\qquad \max_{x \in X} \ u(x).$$

From intermediate micro, choice is not affected by monotonic transformations, if $\psi : \mathbb{R} \to \mathbb{R}$ and $r > s$ implies that $\psi(r) > \psi(s)$, we can set $v(x) = \psi(u(x))$, and x^* is also the solution to

$$(2) \qquad\qquad \max_{x \in X} \ v(x).$$

Example: with $X = \{x_1, \ldots, x_4\}$, suppose that I like x_3 better than all of them, that x_1 and x_4 are tied, and both of those are better than x_2. One utility function describing this is

$$(3) \qquad u(x_2) = 1, \ u(x_1) = u(x_4) = 3, \ \text{and} \ u(x_3) = 4.$$

The idea behind this particular utility function is

$$(4) \qquad u(x) = \#\{y : \text{ I like } x \text{ at least as much as I like } y \}.$$

But, any monotonic transformation of $u(\cdot)$ would do just as well, for example $v(x) = \sqrt{x}$.

Question: What behavioral assumptions are we making by interpreting "purposeful behavior" as utility maximizing behavior? Here is a summary of one answer.

1. Given a utility function $u : X \to \mathbb{R}$, define $x \succeq_u y$ if $u(x) \geq u(y)$. This is read "x is at least as prefered as y." $\succeq_u$ has the following properties.
   a. Completeness, for all $x, y \in X$, either $x \succeq_u y$ or $y \succeq_u x$ (or both).
   b. Transitivity: for all $x, y, z \in X$, if $x \succeq_u y$ and $y \succeq_u z$, then $x \succeq_u z$.
2. Given a complete and transitive preference relation $\succeq$ on X, there is a utility function $u : X \to \mathbb{R}$ such that $x \succeq y$ if and only if $u(x) \geq u(y)$.

This means that believing that people's choice behavior can be understood as the solution to the problem "$\max_{x \in X} u(x)$" is the same as the solution to the problem "pick the option in X that is prefered." Completeness and transitivity are minimal assumptions for this model to be internally consistent. Details of these arguments are in Chapter 1.B of the "Notes for a Course in Managerial Economics" on the website for the class. The major theoretic result is the equivalence of utility maximization with what is called the *Weak Axiom of Revealed Preference*, aka WARP.

The Weak Axiom says the following: observable behavior is utility maximizing behavior if and only if, faced with any set of options (a choice situation) where x and y are available, if x is chosen, then for any other set of options containing both x and y, we will never see y chosen unless x is also a choice. To put it another way, if x is revealed by choice behavior to be at least as good as y, then y will never be revealed by choice behavior to be strictly better than x.

Often, especially, but not only in the study of firms, we take the utility to be accurately measured by profits or money. This is a separate assumption according well with intuition and evidence in many cases. But it is not always valid, and we will be crucially interested in the many situations and ways in which it fails to be valid. Trying to figure out what rational, purposeful people are doing passes through the latin question, "*Cui bono?*" or "Who benefits?" With the assumption that the people understand their benefits in monetary terms, this can be reduced to the advice from a famous movie, "Follow the money." But there are many decisions where more/less money is traded off against other aspects of life, and we will be looking at representing these using economic models of maximizing behavior.

2. Unconstrained Optimization

For a microeconomist, the central assumption is that people, as decision makers, are doing as well as they can with the knowledge and resources they have. The mathematics of optimization is the key to understanding the implications of this assumption.

2.1. **The General Form.** Throughout, $\mathbf{x} = (x_1, \ldots, x_n)$ will represent the levels of n different decision variables. When studying different situations, the interpretations of the n variables will change. Examples include production levels for both private and

public goods, investment levels, pollution abatement levels, the amount of time spent on projects, and many many others.

The **unconstrained** optimization problems will have the general form

$$(5) \qquad \max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n) \text{ or } \max_{x_1,\ldots,x_n \geq 0} f(x_1,\ldots,x_n).$$

$f(\cdot)$ is called the **objective function**. It represents what the decision maker is trying to optimize. The second variant of the maximization problem represents situations in which negative levels of the decision variables do not represent anything of interest.

Solutions to the problems in (5) are, by assumption, what the decision maker chooses to do. They will be denoted $\mathbf{x}^* = (x_1^*,\ldots,x_n^*)$. They will have economic interpretations.

2.2. Solving Unconstrained Problems.
Solving $\max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n)$ for $\mathbf{x}^*$, will often be found by solving the system of n equations in n unknowns given by

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_1} = 0$$

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_2} = 0$$

$$\vdots \quad = 0$$

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_n} = 0.$$

Since these equations involve first order derivatives, they are called **First Order Conditions (FOCs)**.

Often, but not always, the FOCs just given can also be used solve the problem $\max_{x_1,\ldots,x_n \geq 0} f(x_1,\ldots,x_n)$. There will be important cases in which this is not true. For those cases, we will need to modify the FOCs, and we will cover this modification later.

2.3. Homework problems on unconstrained maximization.
You MUST know how to do the following problems in order to get through this class. You do not need to hand in the following, but solutions will be given. Think of this as a review of some material from your calculus class.

Review A. Give the FOCs and the solutions to the following problems.
1. $\max_x f(x)$ when $f(x) = 9 + 3x - 4x^2$.
2. $\max_x f(x)$ when $f(x) = -[2x^2 + 4(100 - x)^2]$.
3. $\max_x f(x)$ when $f(x) = 250 + 19x - e^x$.
4. $\max_{x \geq 0} f(x)$ when $f(x) = 50\sqrt{x} - 0.8 \cdot x$.
5. $\max_{x \geq 0} f(x)$ when $f(x) = 500\sqrt{x} - 0.01 \cdot x^2$.
6. $\max_{x \geq 0} f(x)$ when $f(x) = 5\log(x) - 0.1 \cdot x$.
7. $\max_{x \geq 0} f(x)$ when $f(x) = 5\log(x) + 3\log(10 - x)$.
8. $\max_{x \geq 0} f(x)$ when $f(x) = 5x - 0.01x^2$.

Review B. Give the FOCs and the solutions to the following problems.
1. $\max_{x_1,x_2} f(x_1, x_2)$ when $f(x_1, x_2) = 29 + 3x_1 + 4x_2 - (3x_1^2 - 2x_1x_2 + 4x_2^2)$.
2. $\max_{x_1,x_2} f(x_1, x_2)$ when $f(x_1, x_2) = -[x_1^2 + 4x_2^2 + 2(100 - (x_1 + x_2))^2]$.
3. $\max_{x_1,x_2 \geq 0} f(x_1, x_2)$ when $f(x_1, x_2) = 5\log(x_1) + 30\log(x_2) - 0.4 \cdot x_1 - x_2$.

4. $\max_{x_1,x_2 \geq 0} f(x_1, x_2)$ when $f(x_1, x_2) = 15x_1 + 30x_2 - 0.4e^{x_1} - 0.1e^{x_2}$.

Review C. A person has an original amount a of a good. By sacrificing x of it, they can produce $y = g(x)$ of another good. The person solves the utility maximization problem $\max_{x \geq 0} u(a - x, g(x))$. Suppose that $u(c, y) = \log(c) + y$ and that $g(x) = x$.

1. Give the FOCs for the maximization problem $\max_{x \geq 0} u(a - x, g(x))$.

2. Solve for x^*.

Review D. Person 1 has an original amount a_1 of a good while person 2 has an original amount a_2 of a good. By sacrificing x_1 and x_2 of it, the two of them can produce $y = g(x_1 + x_2)$ of another good. Suppose that $u_1(c_1, y) = \log(c) + y$ and that $u_2(c_2, y) = \log(c) + y$. This is a very simple representation of the idea of a public good — whatever the level of y that is produced, both people enjoy it. As above, suppose that $g(x) = x$.

1. Give the FOCs for the maximization problem

$$\max_{x_1, x_2 \geq 0} \left[u_1(a_1 - x_1, x_1 + x_2) + u_2(a_2 - x_2, x_1 + x_2) \right].$$

2. Solve for $\mathbf{x}^* = (x_1^*, x_2^*)$.

3. Parametrized Optimization Problems

The objective functions of interest are often of interest because they are **parametrized**, $f = f(\mathbf{x}; \theta)$. This changes the problems to

(6) $\qquad V(\theta) = \max_{x_1, \ldots, x_n} f(x_1, \ldots, x_n; \theta)$ or $V(\theta) = \max_{x_1, \ldots, x_n \geq 0} f(x_1, \ldots, x_n; \theta)$.

3.1. **Three Aspects.** There are three aspects to this.

- First, θ is **not** something that the decision maker can choose, it is something outside of their control. When studying different situations, the interpretations of θ will change. Examples include resources, prices of outputs, prices of inputs, pollution reduction targets, measures of benefits and costs included in calculations. There are many others.
- Second, optimizing behavior now depends on the value of the parameter, and we represent this by $\mathbf{x}^*(\theta) = (x_1^*(\theta), \ldots, x_n^*(\theta))$. We care about how $\mathbf{x}^*(\cdot)$ depends on the parameter θ.
- Third, we now include the parametrized **value** of the decision problem, $V(\theta)$.

To find $V(\theta)$ explicitly, we will solve the problem and "plug the solution back in," that is,

(7) $\qquad\qquad V(\theta) = f(x_1^*(\theta), \ldots, x_n^*(\theta); \theta).$

4

3.2. **Examples from Microeconomics.**
- x is the production level for a firm, p is the market price for the good, $c(x)$ is the cost of producing x, and the problem is
$$V(p) = \max_x \ (px - c(x)),$$
that is, $f(x; p) = px - c(x)$. Here: the price p is the parameter; $x^*(p)$ is the supply that the firm produces when the price is p; $V(p) = px^*(p) - c(x^*(p))$ is the profit function.
- x is the production level for a firm, p is the market price for the good, $c(x, w)$ is the cost of producing x when the price of inputs is w, and the problem is
$$V(p, w) = \max_{x \geq 0} \ (px - c(x, w)),$$
that is, $f(x; p, w) = px - c(x, w)$. Here: the parameter is the vector of prices, (p, w); $x^*(p, w)$ is the supply that the firm produces when the price of the output is p and the price of the inputs is w; $V(p, w)$ is the profit function. In intermediate micro, $x^*(\cdot, \cdot)$ is the supply function expressed as a price of both inputs and outputs. We are now explicitly including the dependence of profits on the price of the output and the price of the inputs.
- x_1, x_2 are the production levels for two goods, p_1, p_2 are the market prices for the two goods, $c(x_1, x_2)$ is the cost of production, and the problem is
$$V(p_1, p_2) = \max_{x_1, x_2} \ [(p_1 x_1 + p_2 x_2) - c(x_1, x_2)].$$
that is, $f(x_1, x_2; p_1, p_2) = (p_1 x_1 + p_2 x_2) - c(x_1, x_2)$. The solution vector, $\mathbf{x}^*(p_1, p_2) = (x_1^*(p_1, p_2), x_2^*(p_1, p_2))$ is the joint supply function for the function, and $V(p_1, p_2) = (p_1 x_1^* + p_2 x_2^*) - c(x_1^*, x_2^*)$ is the profit function.
- $x_1, \ldots, x_n$ are the n inputs into the production of good y, the prices of inputs are $w_1, \ldots, w_n$, $y = g(x_1, \ldots, x_n)$ expresses output using the production function $g(\cdot)$, output is sold at a price p. The parameter is now the price of the output as well as the vector of the prices of the inputs,
$$V(p, w_1, \ldots, w_n) = \max_{x_1, \ldots, x_n} \ pg(x_1, \ldots, x_n) - (w_1 x_1 + \cdots + w_n x_n),$$
that is, $f(x_1, \ldots, x_n; p, w_1, \ldots, w_n) = pg(x_1, \ldots, x_n) - (w_1 x_1 + \cdots + w_n x_n)$. The solution vector $\mathbf{x}^*(p, w_1, \ldots, w_n)$ is the vector of goods the firm produces when the output price is p and the input prices are $w_1, \ldots, w_n$. Plugging this vector of solutions back into the objective function gives the profit function.
- When the firm in the previous example is large enough that their decisions affect the price and the demand function is $p(q)$, then we have the price as a function of the decision variables, $p = p(g(x_1, \ldots, x_n))$. In this case, the problem is
$$V(w_1, \ldots, w_n) = \max_{x_1, \ldots, x_n} \ p(g(x_1, \ldots, x_n)) \cdot g(x_1, \ldots, x_n) - (w_1 x_1 + \cdots + w_n x_n).$$
Here $f(x_1, \ldots, x_n; w_1, \ldots, w_n) = p(g(x_1, \ldots, x_n)) \cdot g(x_1, \ldots, x_n) - (w_1 x_1 + \cdots + w_n x_n)$ and the solution vector is the supply vector as function of the input prices. In the previous example, the price p was a parameter, it was not under

the control of the decision maker. Here the decision maker *does* control p, hence it is not a parameter.

3.3. **Homework problems on parametrized maximization.**

A. Give $x^*(p)$ and $V(p) = \max_{x \geq 0} (px - c(x))$ when $c(x) = \frac{1}{2}x^2$.

B. Give $x^*(p, w)$ and $V(p, w) = \max_{x \geq 0} (px - c(x, w))$ when $c(x, w) = w(e^x - 1)$.

C. Give $\mathbf{x}^*(p_1, p_2)$ and $V(p_1, p_2) = \max_{x_1, x_2} (p_1 x_1 + p_2 x_2) - c(x_1, x_2)$ when $c(x_1, x_2) = x_1^2 x_2^3$.

D. Give $\mathbf{x}^*(w_1, \ldots, w_n)$ and $V(p, w_1, \ldots, w_n) = \max_{x_1, \ldots, x_n} pg(x_1, \ldots, x_n) - (w_1 x_1 + \cdots + w_n x_n)$ when $g(x_1, \ldots, x_n) = \Pi_{i=1}^n x_i^{\alpha_i}$ where each $\alpha_i > 0$ and $\sum_i \alpha_i < 1$.

E. Give $x^*(a, \beta)$ and $V(a, \beta) = \max_{x \geq 0} [\beta \log(a - x) + x]$. Assume $a > 0$ and $\beta > 0$.

F. Let $V(\alpha, \beta) = \alpha B(e) - C(e)/\beta$ where $B(e) = e$ and $C(e) = e^2/2$.
   1. Find, both algebraically and graphically, how $e^*(\cdot, \cdot)$ depends on both α and β.
   2. Suppose now that the decision maker in this problem has an outside option giving them utility $\underline{u}$. Give the set of α and β for which $V(\alpha, \beta) \geq \underline{u}$. depends on both α and β.

4. Comparative Statics with Derivatives

We now turn to determining the dependence of behavior, $\mathbf{x}^*(\theta)$, on θ. In particular, we will often be interested in knowing whether or not the following kind of monotone results hold:

$$\text{if } \theta^\circ > \theta, \text{ then } \mathbf{x}^*(\theta^\circ) > \mathbf{x}^*(\theta); \text{ or}$$
$$\text{if } \theta^\circ > \theta, \text{ then } \mathbf{x}^*(\theta^\circ) < \mathbf{x}^*(\theta).$$

The key part of the answer is, "If the net marginal benefit of an activity increases as θ increases, then $\mathbf{x}^*(\theta^\circ) > \mathbf{x}^*(\theta)$, if the net marginal benefit decreases, then $\mathbf{x}^*(\theta^\circ) < \mathbf{x}^*(\theta)$." There are subtleties, especially when the parameter and the decision take vector form, but this will be the essential intuition in many many contexts.

4.1. **An Example: Geometry and Calculus.** There are complicated routes to this kind of result, and simple routes. The complicated route is to explicitly calculate $\mathbf{x}^*(\theta)$, then explicitly calculate $\partial \mathbf{x}^*/\partial \theta$, and then check if it is positive or negative. This is overkill, but overkill is often useful. When it can be done, you will not only know how far above or below 0 the derivative $\partial \mathbf{x}^*/\partial \theta$ is, you will also be able to tell what the derivative depends on.

We can often answer the simpler, less detailed, monotone questions — i.e. does the optimum go up or down as the parameter goes up — without needing to do all of the hard work. Consider, as a starting point, the one-input/one-output profit maximization problem,

$$(8) \qquad\qquad V(p) = \max_{x \geq 0} [px - c(x)].$$

In the last set of problems, you solved a version of this problem (with $c(x) = \frac{1}{2}x^2$) by finding and then solving the FOCs. We revisit these with notation that keeps the

parameter, p, more firmly in view. Let $f(x; p) = px - c(x)$, the FOCs are $\partial f(x, p)/\partial x = 0$, that is,

$$p - c'(x) = 0.$$

This is a "net marginal benefit equals 0" equation: the marginal benefit of a small increase in x is p; the marginal cost is $c'(x)$. Since $c'(\cdot)$ is (usually) an increasing function, you can solve this problem by graphing the decreasing function $p - c'(x)$ — it crosses 0 from above at the point $x^*(p)$. To answer the monotone questions, we are interested in what happens to this intersection if p increases to $p^\circ > p$.

- The geometry — if you shift a decreasing function upwards, e.g. shift from the curve $p - c'(x)$ to the everywhere higher curve $p^\circ - c'(x)$, the place where it crosses 0 must move to the right. We therefore know that $x^*(p^\circ) > x^*(p)$. In this example, the economics interpretation of the result is that the supply curve of a competitive firm is increasing in the price of the output.
- The calculus — suppose that $x^*(p)$ is the function that satisfies the FOCs for all p, that is, $p - c'(x^*(p)) \equiv 0$. Suppose also that $x^*(\cdot)$ has a derivative. Taking the derivative of the FOCs along the curve $x^*(p)$ **with respect to the parameter** p involves finding

$$\tfrac{d}{dp}(p - c'(x^*(p))),$$

  and this yields (using the chair rule from calculus)

(9) $$1 - c''(x^*(p))\tfrac{dx^*(p)}{dp} = 0, \text{ or } \tfrac{dx^*(p)}{dp} = \tfrac{1}{c''(x^*(p))}.$$

  The assumption that $c'(\cdot)$ is increasing is the assumption that $c''(x) > 0$. The detailed information is $\frac{dx^*(p)}{dp} = \frac{1}{c''(x^*(p))}$, the monotone information $\frac{dx^*(p)}{dp} > 0$.

4.2. **Adding Another Parameter.** Now let us add a second detail to the problem in (8) a bit, replacing it with

(10) $$V(p, w) = \max_{x \geq 0} [px - w \cdot c(x)].$$

The new question is how the optimal behavior depends on w, the price of the input into the productive process. The FOCs are $\partial f(x; p, w)/\partial x = 0$.

- Geometry — the net marginal benefit is the decreasing function $p - wc'(x)$, if w increases to $w^\circ > w$, the decreasing function shifts down to $p - w^\circ c'(x)$ and the intersection, $x^*(p, w)$, moves to the left.
- Calculus — if $p - wc'(x^*(p, w)) \equiv 0$, taking derivatives of both sides **with respect to the parameter** w yields (using the product rule from calculus)

(11) $$-c'(x^*(p, w)) - wc''(x^*(p, w))\tfrac{\partial x^*(p,w)}{\partial w}, \text{ or } \tfrac{\partial x^*(p,w)}{\partial w} = -\tfrac{c'(x^*(p,w))}{wc''(x^*(p,w))}.$$

  The detailed information is the complicated expression for $\frac{\partial x^*(p,w)}{\partial w}$, the monotone information is that $\frac{\partial x^*(p,w)}{\partial w} < 0$ because $c'(x) > 0$ and $c''(x) > 0$.

7

4.3. **Homework problems on comparative statics with derivatives.** It will be helpful for the course if, before you start doing the calculations, you ask yourself if net marginal benefits are increasing or decreasing in the parameter in question.

G. From Kreps, problem 11.2, p. 262, on the basics of a simple cost functions.
H. From Kreps, problem 11.10, p. 265, on multiple process cost functions.
I. From Kreps, problem 12.1, p. 289, on short-run vs. long-run cost functions.
J. Give the FOCs for $V(p, w) = \max_{x \geq 0} (px - wc(x))$ when $c(x) = x^2 + (e^x - 1)$. [If you can solve the FOCs for $x^*(p, w)$ as a function of p and w in terms of known functions, then you have made a mistake.] Show that $\partial x^*(p, w)/\partial p > 0$ and $\partial x^*(p, w)/\partial w < 0$ by checking the conditions discussed just above.
K. Suppose that $u(c, y) = \beta \log(c) + y$, that $g(x) = \sqrt{x + x^{1/3}}$, and give the FOCs for $V(\beta) = \max_{x \geq 0} \beta \log(a - x) + g(x))$. Is $\partial x^*(\beta)/\partial \beta > 0$? Or < 0?
L. Suppose that $u(c, y) = \beta \log(c) + y$, that $g(x) \geq 0$, that $g'(x) > 0$, and that $g''(x) < 0$. Give the FOCs for $V(\beta) = \max_{x \geq 0} \beta > \log(a - x) + g(x))$, $\beta > 0$. Is $\partial x^*(\beta)/\partial \beta > 0$? Or < 0?
M. Suppose that $u(c, y) = \beta \log(c) + y$, that $g(x) \geq 0$, that $g'(x) > 0$, and that $g''(x) < 0$. Give the FOCs for $V(\beta, \gamma) = \max_{x \geq 0} \beta \log(a - x) + \gamma g(x))$, $\beta, \gamma > 0$. Is $\partial x^*(\beta, \gamma)/\partial \gamma > 0$? Or < 0?
N. A timing problem. Suppose that by waiting until $t \geq 0$, benefits (in dollars) are $B(t)$ where $B'(t) > 0$ and $B(\cdot)$ is "ess-shaped," that is, it has a region of convexity, i.e. for $0 < t < \underline{t}$), $B''(t) > 0$, and then a region of concavity, i.e. for $\underline{t} < t < \infty$), $B''(t) < 0$. The parametrized problem for maximizing discounted present value is $V(r) = \max_{t \geq 0} B(t)e^{-rt}$ where $r > 0$ is the interest rate/opportunity cost of capital. Give and interpret the FOCs and give the sign for $dt^*(r)/dr$.

5. Monotone Comparative Statics: Super- and Sub-modularity

Above, we took the derivatives of FOCs with respect to parameters to find the monotone results. We are now going to replace the derivative-based analysis with something that is simultaneously easier and more general. This is possible because it often happens that the more general approach is simpler — it lets you focus on the essentials and ignore the complicated details.

There is an 'entry cost' for this kind of analysis, learning to manipulate inequalities. This looks more difficult than it is. You should find that the derivative analysis above often provides an easy guide.

5.1. **The Supermodular Setting.** We start with a set $X \subset \mathbb{R}$, a set $\Theta \subset \mathbb{R}$, and a function $f : X \times \Theta \to \mathbb{R}$. For any $\theta \in \Theta$, let $x^*(\theta)$ be the solution (or set of solutions) to the stripped-down problem

$$(12) \qquad \max_{x \in X} f(x, \theta).$$

We are interested in the comparison of $x^*(\theta^\circ)$ and $x^*(\theta)$ when $\theta^\circ > \theta$.

For $x \in X \subset \mathbb{R}$, $\theta \in \Theta \subset \mathbb{R}$, a function $f : X \times \Theta \to \mathbb{R}$ **supermodular** if for all $\theta^\circ > \theta$ and all $x^\circ > x$,

$$(13) \qquad f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \geq f(x^\circ, \theta) - f(x, \theta),$$

and it is **strictly supermodular** if the inequality is always strict.

Another name for supermodularity is **increasing differences** — the difference $f(x^\circ, \theta) - f(x, \theta)$ is higher at higher values of θ, that is, the difference increases as θ increases.

The function is **submodular** if for all $\theta^\circ > \theta$ and all $x^\circ > x$,

(14) $$f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \leq f(x^\circ, \theta) - f(x, \theta),$$

and it is **strictly submodular** if the inequalities are strict.

Another name for submodularity is **decreasing differences**.

With some qualifications (having to do with the possibility of multiple optima), the essential result is the following.

Super-modularity and comparative statics. If $f(\cdot, \cdot)$ is supermodular and $\theta^\circ > \theta$, then $x^*(\theta^\circ) \geq x^*(\theta)$.

Proof. For now we limit ourselves to the case that there is only one optimal x at any given θ. In other words, we are giving the argument for the super-modularity and comparative statics result under the additional assumption that $x^*(\theta)$ contains at most one element for each θ.

Suppose that $f(\cdot, \cdot)$ is supermodular that $\theta^\circ > \theta$, that x^* is optimal at the lower value, θ, and that x is some point less that x^*. Because x^* is optimal at θ, we know that $f(x^*, \theta) > f(x, \theta)$, that is,

$$f(x^*, \theta) - f(x, \theta) > 0.$$

Because $f(\cdot, \cdot)$ is supermodular, equation (13) holds, i.e. $f(x^*, \theta^\circ) - f(x, \theta^\circ) \geq f(x^*, \theta) - f(x, \theta)$. Therefore

$$f(x^*, \theta^\circ) - f(x, \theta^\circ) > 0.$$

This means that any $x < x^*$ cannot be optimal at θ°. From this, we can conclude that if there is an optimum at the higher value of the parameter, θ°, then that optimum must be greater than or equal to x^*. □

Sub-modularity and comparative statics. If $f(\cdot, \cdot)$ is submodular and $\theta^\circ > \theta$, then $x^*(\theta^\circ) \leq x^*(\theta)$.

The argument is almost the same.

5.2. **Some Examples.** We begin with a general observation that will make checking super/submodularity easier.

5.2.1. *On Cancellations.* Return to the single-input/single output competitive firm, the one that solves

$$\max_{x \in X} \ f(x, (p, w)) = px - wc(x)$$

where $X \subset \mathbb{R}_+$ is the set of possible production levels. We will show that $f(\cdot, \cdot)$ is supermodular in x and p and submodular in x and w.

Supermodularity. Pick $p^\circ > p$ and $x^\circ > x$. To show supermodularity we must show that

$$f(x^\circ, (p^\circ, w)) - f(x, (p^\circ, w)) \geq f(x^\circ, (p, w)) - f(x, (p, w)), \text{ that is}$$
$$[p^\circ x^\circ - wc(x^\circ)] - [p^\circ x - wc(x)] \geq [px^\circ - wc(x^\circ)] - [px - wc(x)].$$

The $wc(x^\circ)$ and the $wc(x)$ terms appear on both sides and cancel.

This kind of cancellation happens all the time. Pay attention to it.

After the cancellation, all that we need to check is

$$[p^\circ x^\circ - p^\circ x] \geq [px^\circ - px], \text{ that is}$$
$$p^\circ[x^\circ - x] \geq p[x^\circ - x].$$

We know that $[x^\circ - x] > 0$ and we know that $p^\circ > p$, so $f(\cdot, \cdot)$ is strictly supermodular in x and p.

Submodularity. Pick $w^\circ > w$ and $x^\circ > x$. To show submodularity we must show that

$$f(x^\circ, (p, w^\circ)) - f(x, (p, w^\circ)) \leq f(x^\circ, (p, w)) - f(x, (p, w)), \text{ that is}$$
$$[px^\circ - w^\circ c(x^\circ)] - [px - w^\circ c(x)] \leq [px^\circ - wc(x^\circ)] - [px - wc(x)].$$

The px° and the px terms appear on both sides and cancel. *The cancellation happened again!* All that we need to check that

$$[w^\circ c(x) - w^\circ c(x^\circ)] \leq [wc(x) - wc(x^\circ)] \text{ that is}$$
$$w^\circ[c(x) - c(x^\circ)] \leq w[c(x) - c(x^\circ)].$$

Because $x < x^\circ$, $c(x) \leq c(x^\circ)$. If they are equal then we have the requisite inequality holding as an equality, if they are unequal, then the requisite inequality holds strictly.

More generally, suppose that $f(x, \theta) = g(x, \theta) + h(x) + m(\theta)$ where $h(\cdot)$ and $m(\cdot)$ are arbitrary functions. To check the inequalities for checking supermodularity of $f(\cdot, \cdot)$, we can ignore the $h(x)$ and the $m(\theta)$ terms — they will cancel.

The general lesson: you only need to pay attention to terms that include the action, x, and the parameter θ. To see why, pick $x^\circ > x$ and $\theta^\circ > \theta$. We have

$$f(x^\circ, \theta^\circ) - f(x, \theta^\circ) = [g(x^\circ, \theta^\circ) - g(x, \theta^\circ)] +$$
$$[h(x^\circ) - h(x)] + [m(\theta^\circ) - m(\theta^\circ)].$$

We also have

$$f(x^\circ, \theta) - f(x, \theta) = [g(x^\circ, \theta) - g(x, \theta)] +$$
$$[h(x^\circ) - h(x)] + [m(\theta) - m(\theta)].$$

To check that $[f(x^\circ, \theta^\circ) - f(x, \theta^\circ)] \geq [f(x^\circ, \theta) - f(x, \theta)]$, note that the $[h(x^\circ) - h(x)] + [m(\theta^\circ) - m(\theta^\circ)]$ and the $[h(x^\circ) - h(x)] + [m(\theta) - m(\theta)]$ terms cancel.

Returning to the example above, to show that $f = px - wc(x)$ is supermodular in p and x, we only need check that $g = px$ is supermodular. The inequality

$$[p^\circ x^\circ - p^\circ x] > [px^\circ - px]$$

is immediate because $x° > x$ and $p° > p$. To check submodularity in w and x, we only need check that $g = -wc(x)$ is submodular. If you prefer, it is equivalent to check that $wc(x)$ is supermodular in w and x because multiplying the inequalities in (13), those that define supermodularity, by -1 changes their direction, giving the inequalities that define submodularity, (14). Anyhow, For $w° > w$ and $x° > x$, this involves checking

$$[w°c(x°) - w°c(x)] \geq [wc(x°) - wc(x)],$$

that is, $w°[c(x°) - c(x)] \geq w[c(x°) - c(x)]$ which is immediate.

5.2.2. *On Monopoly and Monopsony.* When markets break down, it almost aways has an inimical effect on society. When there is only one person/organization on the supply side of the market, we have a **monopoly**, when there is only one person/organization on the demand side of the market, we have a **monopsony**. Both forms of market breakdown have happend at different points in history, and the consequences have varied from merely bad to outright evil. Let us give the bloodless analysis first, then cite some examples.

Monopoly. The demand curve is $p(q)$, the cost curve is $C(q)$, the monopolist solves the problem

(15) $$\max_{q \geq 0} [qp(q) - C(q)].$$

When q' is sold at the price $p' = p(q')$, consumer surplus is

$$S(q) := \int_0^q [p(q) - p']\, dq.$$

Consumer surplus is an increasing function, for $q° > q$, $S(q°) > S(q)$.

Society's problem is

(16) $$\max_{q \geq 0} [(qp(q) + S(q)) - C(q)].$$

The problems (15) and (16) can be put together by setting $f(q,\theta) = [(qp(q) + \theta S(q)) - C(q)]$ and setting $\Theta = \{0,1\}$.

- The problem (15) is $\max_{q \geq 0} f(q,0)$, while
- the problem (16) is $\max_{q \geq 0} f(q,1)$.

The function $f(q,\theta)$ is strictly supermodular in q and θ — to check, we need only look at the term $\theta S(q)$, pick $q° > q$ and $1 = \theta° > \theta = 0$, and check that

$$[1S(q°) - 1S(q)] > [0S(q°) - 0S(q)],$$

which holds because consumer surplus is an increasing function. This means that the monopolist produces less than the amount that maximizes the sum of producer and consumer surplus. By producing more, society is made better off, and in moving to the higher quantity, that which maximizes society's welfare, the winners i.e. the consumers, can compensate the losers i.e. the owner(s) of the monopoly.

Examples. Lachlan Macquarie breaking the English army's monopoly on the medium of exchange in Australia, "rum," meant that the economy could move from barter to market-mediated exchange. In the late 19'th and early 20'th century, the Northern Securities Co. had a railroad monopoly on freight from northern mid-west farms to cities in the U.S. In the U.S. broadband is far slower and costs far more than in other

countries, a result of low levels of competition allowed by a number of Federal and State laws. When you are the monopoly supplier of arms used in war and you price as a monopolist, you are *war profiteering.* This was defined, and prosecuted, as treason against the U.S. in WWII. It was not prosecuted during the invasion of Iraq.

Monopsony. The labor supply curve is an increasing function $w(q)$. Revenue for the single firm buying labor in the local market is $R(q)$. The easy case is $R(q) = pf(q)$ where $f(\cdot)$ is the production function, we expect $f'(\cdot)$ to be a positive, decreasing function, and we expect $R(\cdot)$ to have the same properties. The monopsonist solves the problem

$$(17) \qquad\qquad \max_{q \geq 0} \left[R(q) - qw(q) \right].$$

When a quantity of labor q' is hired at wages $w' = w(q')$, the surplus of the workers is $S(q') := \int_0^{q'} [w' - w(q)] \, dq$, an increasing function. Society problem is

$$(18) \qquad\qquad \max_{q \geq 0} \left[(R(q) + S(q)) - qw(q) \right].$$

Set $f(q, \theta) = (R(q) + \theta S(q)) - qw(q)$, for $\theta = 0$ we have the monopsonist's problem, for $\theta = 1$, we have society's problem, check that $f(\cdot, \cdot)$ is strictly supermodular in q and θ, which means that the monopsonist decreases wages relative to the social optimum, and that it is possible to raise monopsony wages and have the winners compensate the losers.

  Examples. Company towns. Suppliers, e.g. of fighter jet engines, who only have one buyer.

5.3. **The Relation to the Derivative Arguments.** Though the argument for the super-modularity and comparative statics result made no use of derivatives, and may therefore feel unfamiliar, it is related to the "derivative of the FOCs" work you did above.

  • Let $x^\circ = x + dx$ for some small, positive dx. Dividing both sides of equation (13) by dx yields

$$(19) \qquad\qquad \frac{f(x+dx,\theta^\circ) - f(x,\theta^\circ)}{dx} \geq \frac{f(x+dx,\theta) - f(x,\theta)}{dx}.$$

  • Taking $dx \downarrow 0$ (as you did in your calculus classes) yields

$$(20) \qquad\qquad \frac{\partial f(x,\theta^\circ)}{\partial x} \geq \frac{\partial f(x,\theta)}{\partial x},$$

    that is, higher values of the parameter θ shift the marginal net benefit curve upward, hence shift where it crosses 0 to the right.
  • The previous can be re-written in terms of the cross-partial derivatives of $f(\cdot, \cdot)$. Let $\theta^\circ = \theta + d\theta$ for a small positive $d\theta$ and then send $d\theta \downarrow 0$ (as in your calculus classes). This yields

$$(21) \qquad\qquad \frac{\partial f(x,\theta+d\theta)}{\partial x} - \frac{\partial f(x,\theta)}{\partial x} \geq 0$$

$$(22) \qquad\qquad \frac{\frac{\partial f(x,\theta+d\theta)}{\partial x} - \frac{\partial f(x,\theta)}{\partial x}}{d\theta} \geq 0$$

$$(23) \qquad\qquad \frac{\partial^2 f(x,\theta)}{\partial x \partial \theta} \geq 0.$$

When the function $f(\cdot, \cdot)$ is differentiable, the following result often makes it easy to check for supermodularity.

Super-modularity for differentiable functions. If $f(\cdot, \cdot)$ is twice continuously differentiable and for all (x, θ), $\frac{\partial^2 f(x,\theta)}{\partial x \partial \theta} \geq 0$, then $f(\cdot, \cdot)$ is supermodular.

5.4. **Homework Problems on Supermodular Comparative Statics.**

O. A biotech firm spends $x \geq 0$ researching a cure for a rare condition (for example, one covered by the Orphan Drug Act), its expected benefits are $B_1(x)$, the social benefits not capturable by the firm are $B_2(x)$, and both are increasing functions.
   1. Show that the optimal x is larger than the one the firm would choose.
   2. Show that allowing the firm to capture more of the social benefits (e.g. by giving longer patents or subsidizing the research), governments can increase the x that the firm chooses.
P. One part of the business model of a consulting company is to hire bright young men and women who have finished their undergraduate degrees and to work them long hours for pay that is low relative to the profits they generate for the company. The youngsters are willing to put up with this because the consulting company provides them with a great deal of training and experience, all acquired over the course of the, say, three to five years that it takes for them to burn out, to start to look for a job allowing a better balance of the personal and professional. The value of the training that the consulting company provides is at least partly recouped by the youngsters in the form of higher compensation at their new jobs. Show that the consulting company is probably providing an inefficiently low degree of training.
Q. When one looks at statistics measuring the competence with which firms are run, after adjusting for the industry, one finds a weak effect in favor of firms with female CEO's, and a much stronger effect in favor of larger firms. A good part of this is that well-run firms are the ones that succeed and grow, so when you look at firms presently in existence, the well-run ones are larger. In this problem, you are going to investigate a different advantage of being large, the decreasing average cost aspect of simple inventory systems. Decreasing average costs sometimes go by the name of economies of scale, and economies of scale are a crucial determinant of the horizontal boundary of a firm. In this problem, you will find a power law relating size to costs.

   Your firm needs Y units of, say, high grade cutting oil per year. Each time you order, you order an amount Q at an ordering cost of $F + pQ$, where F is the fixed cost of making an order (e.g. you wouldn't want just anybody to be able to write checks on the corporate account and such sytems are costly to implement), and p is the per unit cost of the cutting oil. This means that your yearly cost of ordering is $\frac{Y}{Q} \cdot (F + pQ)$ because $\frac{Y}{Q}$ is the number of orders per year of size Q that you make to fill a need of size Y.

   Storing anything is expensive, and the costs include insurance, the opportunity costs of the space it takes up, the costs of keeping track of what you actually have, and so on. We suppose that these stockage costs are s per unit stored. Computerized records and practices like bar-coding have substantially reduced s

over the last decades. Thus, when you order Q and draw it down at a rate of Y per year, over the course of the cycle that lasts Q/Y of a year, until you must re-order, you store, on average $Q/2$ units. This incurs a per year cost of $s \cdot \frac{Q}{2}$. Putting this together, the yearly cost of running an inventory system to keep you in cutting oil is

$$(24) \qquad C(Y) = \min_Q \; \left[\frac{Y}{Q} \cdot (F + pQ) + s \cdot \frac{Q}{2} \right],$$

and the solution is $Q^*(Y, F, p, s)$.

1. Without actually solving the problem in equation (24), find out whether Q^* depends positively or negatively on the following variables, and explain, in each case, why your answers makes sense: Y; F; p; and s.
2. Now explicitly find the optimal tradeoff between fixed costs and storage costs to solve for $Q^*(Y, F, p, s)$ and $C(Y)$.
3. Find the marginal cost of an increase in Y. Verify that the average cost, $AC(Y)$, is decreasing and explain how your result about the marginal cost implies that this must be true.
4. With the advent and then lowering expenses of computerized inventory and accounting systems, the costs F and s have both been decreasing. Does this increase or decrease the advantage of being large?

R. This problem provides an explanation for the observation that people put in more effort in more challenging environments. Suppose that effort, $e \geq 0$, always costs $C(e)$ where $C(\cdot)$ is an increasing function. In situation h, the benefits of effort are given by the increasing function $B_h(e)$, in situation u, they are given by the increasing function $B_u(e)$. The function $B_h(\cdot)$ comes very close it its maximum at very low levels of e, while the function $B_u(\cdot)$ has a larger slope, that is, for all $e' > e$, $[B_u(e') - B_u(e)] > [B_h(e') - B_h(e)]$.

1. Compare $e^*(h)$ to $e^*(u)$.
2. How would your answer change if $C(\cdot)$ had regions of decrease?

6. Quasi-supermodularity

Lest you think that everything has become easy, let us consider what happens to a monopolist's supply after the demand curve shifts inwards or outwards by some factor $\theta > 0$. If the demand curve of the monopolist shifts from $p(q)$ to $\theta \cdot p(q)$ where $\theta > 0$, consider the problems

$$(25) \qquad V(\theta) = \max_q \; \pi(q, \theta) = [q\theta p(q) - wc(q)].$$

If we knew that $\pi(\cdot, \cdot)$ had increasing differences in q and θ, we would know that outward expansions of the demand curve would increase suppy, but this does not hold here — $Rev(q) = p \cdot p(q)$ increases and then decreases for all reasonable demand functions. However, $\pi(q, \theta)$ is **quasi-supermodular** in q and θ, and it is this that we use in our arguments that the optima are an increasing function of θ.

For $x \in X \subset \mathbb{R}$, $\theta \in \Theta \subset \mathbb{R}$, a function $f : X \times \Theta \to \mathbb{R}$ **quasi-supermodular** if for all $\theta^\circ > \theta$ and all $x^\circ > x$,

(26) $\qquad$ if $f(x^\circ, \theta) - f(x, \theta) > 0$, then $f(x^\circ, \theta^\circ) - f(x, \theta^\circ) > 0$, and

(27) $\qquad$ if $f(x^\circ, \theta) - f(x, \theta) \geq 0$, then $f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \geq 0$.

You should check that every supermodular function is quasi-supermodular. Further, the argument that we gave for supermodular functions having increasing optima actually only used quasi-supermodularity.

S. Show that $\pi(q, \theta)$ given above is quasi-supermodular but not supermodular.
T. Give complete arguments for the following.
1. A supermodular function is quasi-supermodular.
2. If $f(x, \theta)$ is quasi-supermodular and $\varphi(\cdot)$ is an increasing function, then $h(x, \theta) = \varphi(f(x, \theta))$ is quasi-supermodular.
3. There are quasi-supermodular functions that are not supermodular.
4. If $f(\cdot, \cdot)$ is quasi-supermodular and $\theta^\circ > \theta$, then $x^*(\theta^\circ) \geq x^*(\theta)$. Assume that there is at most one optimizing solution at any θ.

7. Constrained Maximization and Lagrangeans

For a microeconomist, the central assumption is that people, as decision makers, are doing as well as they can with the knowledge and resources they have. We now turn to incorporating the constrainst on resources. We will discuss constraints on knowledge later in the semester.

7.1. **The General Forms.** The **constrained** optimization problems will come in one of four forms:

- one constraint,

(28) $\qquad V(b) = \max\limits_{x_1, \dots, x_n} f(x_1, \dots, x_n)$ subject to $g(x_1, \dots, x_n) \leq b$,

- m constraints,

(29) $\qquad V(b_1, \dots, b_m) = \max\limits_{x_1, \dots, x_n} f(x_1, \dots, x_n)$ subject to

(30) $\qquad\qquad\qquad\qquad\qquad g_1(x_1, \dots, x_n) \leq b_1$

$$\vdots$$

(31) $\qquad\qquad\qquad\qquad\qquad g_m(x_1, \dots, x_n) \leq b_m$

- one constraint plus non-negativity,

(32) $\quad V(b) = \max\limits_{x_1, \dots, x_n} f(x_1, \dots, x_n)$ subject to $g(x_1, \dots, x_n) \leq b$ and $x_1, \dots, x_n \geq 0$,

- m constraints plus non-negativity,

$$(33) \qquad V(b_1, \ldots, b_m) = \max_{x_1, \ldots, x_n} f(x_1, \ldots, x_n) \text{ subject to}$$

$$(34) \qquad g_1(x_1, \ldots, x_n) \le b_1$$

$$\vdots$$

$$(35) \qquad g_m(x_1, \ldots, x_n) \le b_m$$

$$(36) \qquad x_1, \ldots, x_n \ge 0.$$

7.2. Solving Constrained Problems.

Solving $\max_{x_1, \ldots, x_n} f(x_1, \ldots, x_n)$ subject $g(x_1, \ldots, x_n) \le b$ for $\mathbf{x}^*$, will often be found in two steps. First, one writes out the Lagrangean function,

$$(37) \qquad L(x_1, \ldots, x_n; \lambda) = f(x_1, \ldots, x_n) + \lambda(b - g(x_1, \ldots, x_n)).$$

Then one solves the system of $n+1$ equations in $n+1$ unknowns given by

$$(38) \qquad \frac{\partial L(x_1, \ldots, x_n; \lambda)}{\partial x_1} = 0$$

$$(39) \qquad \frac{\partial L(x_1, \ldots, x_n; \lambda)}{\partial x_2} = 0$$

$$(40) \qquad \vdots \quad = 0$$

$$(41) \qquad \frac{\partial L(x_1, \ldots, x_n; \lambda)}{\partial x_n} = 0$$

$$(42) \qquad \frac{\partial L(x_1, \ldots, x_n; \lambda)}{\partial \lambda} = 0.$$

Since these equations involve first order derivatives, they are called **First Order Conditions (FOCs)**.

When there are m constraints, this becomes $n+m$ equations in $n+m$ unknowns. Incorporating the non-negativity constraints will be covered later.

We will mostly work with functions f and g for which the following is true: if $(\mathbf{x}^*, \lambda^*)$ solves the FOCs, then $g(\mathbf{x}^*) \le b$, $f(\mathbf{x}^*) \ge f(\mathbf{x}')$ for any $\mathbf{x}'$ satisfying $g(\mathbf{x}') \le b$, and $\lambda^* = \partial V(b)/\partial b$.

7.3. Homework Problems on Constrained Optimization.

U. Let $f(x_1, x_2) = 1,500 - [(x_1 - 100)^2 + (x_2 - 100)^2]$, let $g(x_1, x_2) = x_1 + x_2$ and $b = 40$.
  1. Write out the Lagrangean for the problem $\max f(x_1, x_2)$ subject to $g(x_1, x_2) \le b$.
  2. Write out the FOCs for the Lagrangean.
  3. Solve the FOCs for $\mathbf{x}^*$ and $\partial V(b)/\partial b$ at $b = 40$.

V. Let $f(x, y) = 210 \cdot y - 0.01x$ subject to $g(x, y) \le 0$ where $g(x, y) = y - \sqrt{x}$.
  1. Write out the Lagrangean for the problem $\max f(x, y)$ subject to $g(x, y) \le 0$.
  2. Write out the FOCs for the Lagrangean.
  3. Solve the FOCs for (x^*, y^*) and $\partial V(b)/\partial b$ at $b = 0$.

W. Solve $\max_{x_1, x_2 \ge 0} [0.3 \log(x_1) + 0.4 \log(x_2)]$ subject to $11x_1 + 2x_2 \le 104$ and give the derivative of the value function at $b = 104$.

X. Solve $\min_{x_1, x_2, x_3 \ge 0} [7x_1^2 + 9x_2^2 + x_3^2]$ subject to $(x_1 + x_2 + x_3) \ge X$ as a function of X and give $\partial V(X)/\partial X$.

8. Expected Utility and Diversification

We capture the idea of choice in the presence of risk as the choice between different random variables. We need to associated these objects of choice with some utility numbers. Expected utility theory supposes that you like random variable X better than Y if

$$E\,u(X) > E\,u(Y) \tag{43}$$

where $u(\cdot)$ is known as the **expected utility function**.

8.1. Risk Aversion and Concavity.
"Degenerate" random variables are constant, i.e. they have variance 0. Risk aversion is prefering the mean of a random variable for sure to the random variable, that is,

$$u(E\,X) > E\,u(X). \tag{44}$$

This is where concavity and Jensen's inequality enter the analysis.

8.2. Homework Problems on Expected Utility Theorem.
Y. From Kreps, problems 18.1 and 2, p. 424-5, on choice between random variables.
Z. From Kreps, problems 18.3 and 18.4, p. 425-6, on initial wealth levels and partial insurance.

Kreps, problems 18.6, p. 427, is a start on the benefits of diversification. The next set of problems will spend a good bit of time on this idea and related topics.