

Assignment #1 for **Environmental and Resource Economics**
Economics 359M, Spring 2017

Due date: Wednesday, February 7, 2017

Readings: Chapters 1, 2, and 18 in

Kolstad. *Environmental Economics, 2nd ed.* OUP.

A. Sandmo. The early history of environmental economics. *Review of Environmental Economics and Policy*, 9(2):1–21, 2015.

If you have already seen a good treatment of the mathematics of constrained optimization, then you can go straight to the problems. If not, these notes, or some other source (e.g. your calculus textbook, a calculus-based microeconomics textbook) are part of your reading assignment.

CONTENTS

1. Unconstrained Optimization	2
1.1. The General Form	2
1.2. Solving Unconstrained Problems	2
1.3. Homework problems on unconstrained maximization	2
2. Parametrized Optimization Problems	4
2.1. Three Aspects	4
2.2. Examples from Microeconomics	5
2.3. Homework problems on parametrized maximization	6
3. Comparative Statics with Derivatives	6
3.1. An Example: Geometry and Calculus	7
3.2. Adding Another Parameter	8
3.3. Homework problems on comparative statics with derivatives	8
4. Monotone Comparative Statics: Super- and Sub-modularity	8
4.1. The Supermodular Setting	9
4.2. Some Examples	10
4.3. The Relation to the Derivative Arguments	13
4.4. Homework Problems on Supermodular Comparative Statics	14
5. Constrained Maximization and Lagrangeans	15
5.1. The General Forms	15
5.2. Solving Unconstrained Problems	16
5.3. Homework Problems Using Lagrangeans	16

1. 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.

1.1. The General Form. Throughout, $\mathbf{x} = (x_1, \dots, x_n)$ 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

$$\max_{x_1, \dots, x_n} f(x_1, \dots, x_n) \text{ or } \max_{x_1, \dots, x_n \geq 0} f(x_1, \dots, x_n). \quad (1)$$

$f(\cdot)$ 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 (1) are, by assumption, what the decision maker chooses to do. They will be denoted $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$. They will have economic interpretations.

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

$$\begin{aligned} \frac{\partial f(x_1, \dots, x_n)}{\partial x_1} &= 0 \\ \frac{\partial f(x_1, \dots, x_n)}{\partial x_2} &= 0 \\ &\vdots \\ \frac{\partial f(x_1, \dots, x_n)}{\partial x_n} &= 0. \end{aligned}$$

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 $\max_{x_1, \dots, x_n \geq 0} f(x_1, \dots, x_n)$. 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.

1.3. Homework problems on unconstrained maximization.

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 $f(x) = 5x - 0.01x^2$.
- 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 $f(x_1, x_2) = 15x_1 + 30x_2 - 0.4e^{x_1} - 0.1e^{x_2}$.
- 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^* .
- D. Neonicotinoids are applied to roughly $170 \cdot 10^6$ acres per year in the US, mostly through the treatment of seeds, and they now account for approximately 1/4 of worldwide pesticide sales and 4/5 of the seed treatment. The initial success, after their introduction in the mid-1980's was due to the lack of pesticide resistance in the targets, a variety of advantages over previous generations of pesticides, and an assumption of lower risk for those using the pesticide and those consuming trace amounts. Since that time, they have become widespread in the environment and a number of studies have demonstrated their synergistic and deadly effects with other stressors on pollinating species, especially bees. Suppose that the opportunity and other costs of reducing neonicotinoids by an amount of R is given by $C(R)$ and the benefits of their reduction by $B(R)$.
1. If $C(R) = 100R + \frac{1}{20,000}R^2$ and $B(R)$ is given by $300R$, what is the optimal amount of reduction?
 2. How does the answer change if $C(R)$ is changed to $100R + \frac{1}{20,000}R^2 + C$?
 3. How does the answer change if $B(R)$ is changed to $100R + \frac{1}{20,000}R^2 - B$?
 4. How does the answer change if $C(R)$ is changed to $\alpha \cdot (100R + \frac{1}{20,000}R^2)$?
 5. How does the answer change if $B(R)$ is changed to $\beta \cdot (300R)$?
- E. There are I firms currently emitting quantities q_i of endocrine disruptors (chlorinated pesticides and herbicides, bis-phenolA, di-2-ethylhexyl phthalate, and various surfacants) that are ending up in the local water table. This damages the local

wildlife, especially the fish and amphibians, and it is getting into the local drinking water, which damages the people who drink the water. Simplifying, if the total quantity emitted is $q = \sum_i q_i$, the damages are worth $D(q)$. The cost to firm i of reducing the emissions by an amount r_i is $C_i(r_i)$. The social problem is

$$\max_{r_1, r_2, \dots, r_I} B - D(q - \sum_i r_i) - \sum_i C_i(r_i).$$

Assume throughout this problem that all of the solutions in this problem are strictly positive, that both the function $D(\cdot)$ and the functions $C_i(\cdot)$ have strictly positive derivatives.

1. Give the I different first order conditions for the social problem.
2. Show that the optimal solution involves each firm reducing to a point where their marginal costs are equal. Interpret the quantity that all of the marginal costs are equal to at the optimum.
3. Suppose that there is a price p that the firms must pay to emit a quantity q_i . Profits are $\max_{r_i} [\pi - C_i(r_i) - p(q_i - r_i)]$. Give the first order conditions for

$$\max_{r_i} [\pi - C_i(r_i) - p(q_i - r_i)],$$

and show that they are the same as the first order conditions for

$$\min_{r_i} C_i(r_i) - p \cdot r_i.$$

4. What p must be chosen in the previous problem to have the same answer as in the second problem?

2. 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

$$V(\theta) = \max_{x_1, \dots, x_n} f(x_1, \dots, x_n; \theta) \text{ or } V(\theta) = \max_{x_1, \dots, x_n \geq 0} f(x_1, \dots, x_n; \theta). \quad (2)$$

2.1. **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 $\mathbf{x}^*(\theta) = (x_1^*(\theta), \dots, x_n^*(\theta))$. 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,

$$V(\theta) = f(x_1^*(\theta), \dots, x_n^*(\theta); \theta). \quad (3)$$

2.2. 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

$$V(p_1, p_2) = \max_{x_1, x_2} [(p_1x_1 + p_2x_2) - c(x_1, x_2)].$$

that is, $f(x_1, x_2; p_1, p_2) = (p_1x_1 + p_2x_2) - c(x_1, x_2)$. 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 $V(p_1, p_2) = (p_1x_1^* + p_2x_2^*) - c(x_1^*, x_2^*)$ is the profit function.

- x_1, \dots, x_n are the n inputs into the production of good y , the prices of inputs are w_1, \dots, w_n , $y = g(x_1, \dots, x_n)$ 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,

$$V(p, w_1, \dots, w_n) = \max_{x_1, \dots, x_n} pg(x_1, \dots, x_n) - (w_1x_1 + \dots + w_nx_n),$$

that is, $f(x_1, \dots, x_n; p, w_1, \dots, w_n) = pg(x_1, \dots, x_n) - (w_1x_1 + \dots + w_nx_n)$. The solution vector $\mathbf{x}^*(p, w_1, \dots, w_n)$ is the vector of goods the firm produces when the output price is p and the input prices are w_1, \dots, 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, $p = p(g(x_1, \dots, x_n))$. In this case, the problem is

$$V(w_1, \dots, w_n) = \max_{x_1, \dots, x_n} p(g(x_1, \dots, x_n)) \cdot g(x_1, \dots, x_n) - (w_1x_1 + \dots + w_nx_n).$$

Here $f(x_1, \dots, x_n; w_1, \dots, w_n) = p(g(x_1, \dots, x_n)) \cdot g(x_1, \dots, x_n) - (w_1x_1 + \dots + w_nx_n)$ 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.

2.3. Homework problems on parametrized maximization.

- F. Give $x^*(p)$ and $V(p) = \max_{x \geq 0} (px - c(x))$ when $c(x) = \frac{1}{2}x^2$.
- G. Give $x^*(p, w)$ and $V(p, w) = \max_{x \geq 0} (px - c(x, w))$ when $c(x, w) = w(e^x - 1)$.
- H. Give $\mathbf{x}^*(p_1, p_2)$ and $V(p_1, p_2) = \max_{x_1, x_2} (p_1x_1 + p_2x_2) - c(x_1, x_2)$ when $c(x_1, x_2) = x_1^2x_2^3$.
- I. Give $\mathbf{x}^*(w_1, \dots, w_n)$ and $V(p, w_1, \dots, w_n) = \max_{x_1, \dots, x_n} pg(x_1, \dots, x_n) - (w_1x_1 + \dots + w_nx_n)$ when $g(x_1, \dots, x_n) = \prod_{i=1}^n x_i^{\alpha_i}$ where each $\alpha_i > 0$ and $\sum_i \alpha_i < 1$.
- J. Give $x^*(a, \beta)$ and $V(a, \beta) = \max_{x \geq 0} [\beta \log(a - x) + x]$. Assume $a > 0$ and $\beta > 0$.
- K. Give $\mathbf{x}^*((a_1, a_2), (\beta_1, \beta_2))$ and the value function

$$V((a_1, a_2), (\beta_1, \beta_2)) = \max_{x_1, x_2 \geq 0} [(\beta_1 \log(a_1 - x_1) + x_1) + (\beta_2 \log(a_2 - x_2) + x_2)].$$

3. 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:

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

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 $\mathbf{x}^*(\theta^\circ) < \mathbf{x}^*(\theta)$.”

$\mathbf{x}^*(\theta)$.” There are subtleties, especially when the parameter and the decision take vector form, but this will be the essential intuition in many many contexts.

3.1. 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. 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 without needing to do all of the hard work. Consider, as a starting point, the one-input/one-output profit maximization problem,

$$V(p) = \max_{x \geq 0} [px - c(x)]. \quad (4)$$

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

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

and this yields (using the chain rule from calculus)

$$1 - c''(x^*(p)) \frac{dx^*(p)}{dp} = 0, \text{ or } \frac{dx^*(p)}{dp} = \frac{1}{c''(x^*(p))}. \quad (5)$$

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 $\frac{dx^*(p)}{dp} > 0$.

3.2. Adding Another Parameter. Now let us add a second detail to the problem in (4) a bit, replacing it with

$$V(p, w) = \max_{x \geq 0} [px - w \cdot c(x)]. \quad (6)$$

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)

$$-c'(x^*(p, w)) - wc''(x^*(p, w)) \frac{\partial x^*(p, w)}{\partial w}, \text{ or } \frac{\partial x^*(p, w)}{\partial w} = -\frac{c'(x^*(p, w))}{wc''(x^*(p, w))}. \quad (7)$$

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 $c''(x) > 0$.

3.3. 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.

- L. 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.
- M. Give the FOCs for $V(\beta) = \max_{x \geq 0} \beta \log(a - x) + g(x)$ where $g(x) = \sqrt{x + x^{1/3}}$. Is $\partial x^*(\beta)/\partial \beta > 0$? Or < 0 ?
- N. Suppose that $g(x) \geq 0$, that $g'(x) > 0$, and that $g''(x) < 0$. Give the FOCs for $V(\beta) = \max_{x \geq 0} \beta \cdot \log(a - x) + g(x)$, $\beta > 0$. Is $\partial x^*(\beta)/\partial \beta > 0$? Or < 0 ?
- O. Suppose 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 ?

4. 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.

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

$$\max_{x \in X} f(x, \theta). \quad (8)$$

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 $f : X \times \Theta \rightarrow \mathbb{R}$ **supermodular** if for all $\theta^\circ > \theta$ and all $x^\circ > x$,

$$f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \geq f(x^\circ, \theta) - f(x, \theta), \quad (9)$$

and it is **strictly supermodular** if the inequalities are 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 $x^\circ > x$,

$$f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \leq f(x^\circ, \theta) - f(x, \theta), \quad (10)$$

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 than 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 (9) 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^* . \square

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.

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

4.2.1. 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

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

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

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

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

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

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

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

$$[h(x^\circ) - h(x)] + [m(\theta^\circ) - m(\theta)].$$

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 $x^\circ > x$ and $p^\circ > p$. 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 (9), those that define supermodularity, by -1 changes their direction, giving the inequalities that define submodularity, (10). Anyhow, For $w^\circ > w$ and $x^\circ > x$, this involves checking

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

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

4.2.2. On Monopoly and Monopsony. When markets break down, it almost always 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 happened 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

$$\max_{q \geq 0} [qp(q) - C(q)]. \quad (11)$$

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 $q^\circ > q$, $S(q^\circ) > S(q)$.

Society's problem is

$$\max_{q \geq 0} [(qp(q) + S(q)) - C(q)]. \quad (12)$$

The problems (11) and (12) can be put together by setting $f(q, \theta) = [(qp(q) + \theta S(q)) - C(q)]$ and setting $\Theta = \{0, 1\}$.

- The problem (11) is $\max_{q \geq 0} f(q, 0)$, while
- the problem (12) 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 $q^\circ > q$ and $1 = \theta^\circ > \theta = 0$, and check that

$$[1S(q^\circ) - 1S(q)] > [0S(q^\circ) - 0S(q)],$$

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 enforced by a number of Federal and State laws. War profiteering in WWII and in 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

$$\max_{q \geq 0} [R(q) - qw(q)]. \quad (13)$$

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

$$\max_{q \geq 0} [(R(q) + S(q)) - qw(q)]. \quad (14)$$

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 who only have one buyer.

4.3. 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 (9) by dx yields

$$\frac{f(x+dx, \theta^\circ) - f(x, \theta^\circ)}{dx} \geq \frac{f(x+dx, \theta) - f(x, \theta)}{dx}. \quad (15)$$

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

$$\frac{\partial f(x, \theta^\circ)}{\partial x} \geq \frac{\partial f(x, \theta)}{\partial x}, \quad (16)$$

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

$$\frac{\partial f(x, \theta + d\theta)}{\partial x} - \frac{\partial f(x, \theta)}{\partial x} \geq 0 \quad (17)$$

$$\frac{\frac{\partial f(x, \theta + d\theta)}{\partial x} - \frac{\partial f(x, \theta)}{\partial x}}{d\theta} \geq 0 \quad (18)$$

$$\frac{\partial^2 f(x, \theta)}{\partial x \partial \theta} \geq 0. \quad (19)$$

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.

4.4. Homework Problems on Supermodular Comparative Statics.

- P. Give the arguments for the following statement. If $f(\cdot, \cdot)$ is submodular and $\theta^\circ > \theta$, then $x^*(\theta^\circ) \leq x^*(\theta)$. Assume that there is at most one optimizing solution at any θ .
- Q. A pesticide firm spends $x \geq 0$ researching a new pesticide. Its expected profits are $B(x)$, its expected costs are $C(x)$, and both are increasing functions. There is an additional social cost, $C_a(x)$ that is not borne by the firm, also an increasing function.
1. Show that the firm's optimal x is larger than the one that society would choose.
 2. Show that making the firm pay for more of the damages it causes decreases the x that the firm chooses.
- R. An oil company, i , owns the right to pump as high a flow of oil from their well located over one part of an underground sea of oil. As a function of the flow they choose, f_i , this field make profits year of $\Pi_i(f_i) = pf_i - C_i(f_i, f_{-i})$ where f_{-i} represents the pumping rates of the other oil companies that have rights to pump from this field. The higher the total flow chosen by other companies, the higher are the costs of pumping oil in the future (if you pump too hard, the small openings in the underground rock through which the oil flows begin to collapse). Oil fields often operate under what are called unitization agreements that involve voluntary limitations of pumping rates. Show that without a unitization agreement, or something similar, the flow chosen by i is inefficiently high.
- S. 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.
- T. For $x, t \in [1200, 1900]$, let $f(x, t) = xt$. Since $\partial^2 f / \partial x \partial t = 1$, this function has strictly increasing differences, and since $\partial f(x, t) / \partial x > 0$ for all x, t , $x^*(t) \equiv \{1900\}$. Let $g(x, t) = \log(f(x, t)) = \log(x) + \log(y)$ and note that $\partial^2 g / \partial x \partial t = 0$, strictly increasing differences have disappeared, but $\partial g(x, t) / \partial t > 0$ for all x, t . Let $h(x, t) =$

$\log(g(x, t))$, and $\partial^2 h / \partial x \partial t < 0$, strictly increasing differences have become decreasing differences, but $\partial h(x, t) / \partial x > 0$ for all x, t . The problem $\max_{x \in [1200, 1900]} h(x, t)$ provide an example of strictly submodular function with a constant $x^*(\cdot)$.

U. A function $f : X \times \Theta \rightarrow \mathbb{R}$, $X, \Theta \subset \mathbb{R}$, is said to be **quasi-supermodular** if for all $\theta^\circ > \theta$ in Θ and all $x^\circ > x$ in X , we have

$$\begin{aligned} f(x^\circ, \theta) - f(x, \theta) > 0 &\text{ implies } f(x^\circ, \theta^\circ) - f(x, \theta^\circ) > 0, \text{ and} \\ f(x^\circ, \theta) - f(x, \theta) \geq 0 &\text{ implies } f(x^\circ, \theta^\circ) - f(x, \theta^\circ) \geq 0. \end{aligned}$$

1. Show that a supermodular function is quasi-supermodular.
2. Give a quasi-supermodular function that is not supermodular.
3. Show that 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 .

5. 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 constraint on resources. The constraints on knowledge we will discuss toward the end of the semester.

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

- one constraint,

$$V(b) = \max_{x_1, \dots, x_n} f(x_1, \dots, x_n) \text{ subject to } g(x_1, \dots, x_n) \leq b, \quad (20)$$

- m constraints,

$$V(b_1, \dots, b_m) = \max_{x_1, \dots, x_n} f(x_1, \dots, x_n) \text{ subject to} \quad (21)$$

$$g_1(x_1, \dots, x_n) \leq b_1 \quad (22)$$

\vdots

$$g_m(x_1, \dots, x_n) \leq b_m \quad (23)$$

- one constraint plus non-negativity,

$$V(b) = \max_{x_1, \dots, x_n} f(x_1, \dots, x_n) \text{ subject to } g(x_1, \dots, x_n) \leq b \text{ and } x_1, \dots, x_n \geq 0, \quad (24)$$

- m constraints plus non-negativity,

$$V(b_1, \dots, b_m) = \max_{x_1, \dots, x_n} f(x_1, \dots, x_n) \text{ subject to} \quad (25)$$

$$g_1(x_1, \dots, x_n) \leq b_1 \quad (26)$$

⋮

$$g_m(x_1, \dots, x_n) \leq b_m \quad (27)$$

$$x_1, \dots, x_n \geq 0. \quad (28)$$

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

$$L(x_1, \dots, x_n; \lambda) = f(x_1, \dots, x_n) + \lambda(b - g(x_1, \dots, x_n)). \quad (29)$$

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

$$\frac{\partial L(x_1, \dots, x_n; \lambda)}{\partial x_1} = 0 \quad (30)$$

$$\frac{\partial L(x_1, \dots, x_n; \lambda)}{\partial x_2} = 0 \quad (31)$$

$$\vdots = 0 \quad (32)$$

$$\frac{\partial L(x_1, \dots, x_n; \lambda)}{\partial x_n} = 0 \quad (33)$$

$$\frac{\partial L(x_1, \dots, x_n; \lambda)}{\partial \lambda} = 0. \quad (34)$$

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 $g(\mathbf{x}^*) \leq b$, $f(\mathbf{x}^*) \geq f(\mathbf{x}')$ for any \mathbf{x}' satisfying $g(\mathbf{x}') \leq b$, and $\lambda^* = \partial V(b)/\partial b$.

5.3. Homework Problems Using Lagrangeans.

V. 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 $g(x_1, x_2) \leq b$.
2. Write out the FOCs for the Lagrangean.
3. Solve the FOCs for \mathbf{x}^* and $\partial V(b)/\partial b$ at $b = 40$.

W. Let $f(x, y) = 210 \cdot y - 0.01x$ subject to $g(x, y) \leq 0$ where $g(x, y) = y - \sqrt{x}$.

1. Write out the Lagrangean for the problem $\max f(x, y)$ subject to $g(x, y) \leq 0$.

2. Write out the FOCs for the Lagrangean.
 3. Solve the FOCs for (x^*, y^*) and $\partial V(b)/\partial b$ at $b = 0$.
- X. This problem is about optimal emission reduction. At present, each of I different firms produces a quantity q_i° of NO_x as a by-product of their operations, so that the total quantity produced is $Q^\circ = \sum_i q_i^\circ$. At a cost $c_i(a_i)$, $c'_i > 0$ and $c''_i > 0$, firm i can abate to $q_i^\circ - a_i$. Total abatement is $A = \sum_i a_i$. We wish to reduce emissions from Q° to Q^* , i.e. to abate by an amount $A^* = Q^\circ - Q^*$ in a cost efficient fashion, that is, to solve

$$\min_{a_1, \dots, a_I} \sum_i c_i(a_i) \text{ subject to } \sum_i a_i \geq A^*, a_i \geq 0 \text{ for } i = 1, \dots, I.$$

The first problems ask you to work this with specific functions, the second set of problems asks you to work in more generality. Specifically, for the first set of problems, suppose that $c_i(a_i) = \frac{1}{2}\beta_i(a_i)^2$ where $\beta_1 < \beta_2 < \dots < \beta_I$.

1. Give the Lagrangean for the efficient cost of abatement problem, then take its derivatives and set them equal to 0.
2. Show that the solution is $a_i^* = A^* \frac{1}{\beta_i} \frac{1}{\sum_j 1/\beta_j}$.
3. Suppose now that the NO_x levels, a_i , can be accurately measured and are taxed at a rate τ . Let a_i^τ denote firm i 's profit maximizing amount of abatement when the tax is τ . Show that $a_i^\tau = \tau \frac{1}{\beta_i}$.
4. Find the tax rate $\tau(A^*)$ that achieves total abatement A^* in an efficient manner. Explain **why** $\tau(\cdot)$ should be an increasing function. [If you did the algebra correctly, you will have found an increasing function. I want you to explain why you should have expected the function to be increasing.]
5. Suppose that the NO_x is perfectly mixed, i.e. that the social costs/damage of the total output level Q is given by $C(Q) = \frac{1}{2}\gamma Q^2$. Characterize the optimal tax rate, τ^* , and the optimal abatement, A^* .

For the next set of problems, assume only that $c'_i > 0$ and $c''_i > 0$.

6. Give the Lagrangean for this problem.
7. Assuming that the solution involves each $a_i^* > 0$, show that at the solution, $c'_i(a_i^*) = c'_j(a_j^*)$ for each i, j pair.
8. Suppose now that a_i can be accurately measured and the tax for it is τ . Let a_i^τ denote firm i 's profit maximizing amount of abatement when the tax is τ . Show that profit maximization by the firms will lead to $c'_i(a_i^\tau) = c'_j(a_j^\tau)$ for each i, j pair.
9. Let $\tau(A^*)$ be the tax rate that achieves total abatement A^* in an efficient manner. Show that $\tau(\cdot)$ is increasing.

10. Suppose that the NO_x is perfectly mixed, i.e. that the social costs/damage of the total output level Q is given by $C(Q)$, and that marginal social cost/damage is increasing. Characterize the optimal tax rate, τ , and the optimal abatement A^* .
 11. Explain how a market in tradeable permits can achieve A^* at the same total cost.
- Y. Kolstad, Ch. 18, problems 1 and 2.
Z. Kolstad, Ch. 18, problem 4.