Solutions to Problem Set 2

1. **Gradient is orthogonal to the indifference curve** Let $u : R^N \to R$ be a differentiable function (for concreteness you can think of this as a utility function). Let $I = \{x \in R^N | u(x) = \overline{u}\}$ be a *level set* of the function (in the utility function case, an *indifference curve*). Prove that if $x \in I$ then $\langle x, \nabla u\rangle = 0$ where ∇u is the *gradient* of u.

Answer This proposition is not quite correct as stated. The correct statement is that if x is any point in the (shifted) *tangent hyperplane* to the indifference curve at a point $x_0 \in R^N$ then $\langle \nabla u(x_0), x\rangle = 0$. What is the tangent hyperplane? It is the generalization of a *tangent line* in the case $N = 2$. You can visualize this as a line in the (x_1, x_2) plane that is tangent to the utility function indifference curve at a particular point $x_0 = (x_{1,0}, x_{2,0}) \in R^2$, i.e. the line both a) *touches* the indifference curve at x_0 and b) *has the same slope* as the indifference curve at the point x_0. Notice that a *line* is equivalent to R^1, which is one dimension less than the dimension of the space R^2. In R^N a *hyperplane* is a linear subspace of dimension $N-1$, i.e. one dimension less than the total dimension of the overall space, N. It can be shown that any hyperplane can be represented as the set of all x that are *orthogonal* to some vector $a \in R^N$, i.e. any hyperplane consists of all $x \in R^N$ in the set $H(a)$ given by

$$H(a) = \{x \in R^N | \langle a, x\rangle = 0\}. \tag{1}$$

Since $l(x) = \langle a, x\rangle$ is a linear mapping from R^N to R, the hyperplane is the *null space* of this linear mapping (this is a term you would know if you took linear algebra), and so the hyperplane – the null space — is just the set of all $x \in R^N$ that are orthogonal (perpendicular) to the fixed vector $a \in R^N$.

So a *tangent hyperplane* to a function $u : R^N \to R$ will be a set of points on a hyperplane of R^N that satisfies a) the (shifted) hyperplane *touches* the function $u(x)$ at x_0, and b) the hyperplane has the *same slope* as $u(x)$ at the point x_0 (i.e. it is tangent to $u(x)$ at the point x_0).

How do we write the intuition of tangency down mathematically and show that implies that the gradient of u at x_0 is orthogonal (i.e. perpendicular to) the tangent hyperplane to u at x_0? First let us use a bit more specific notation for the indifference curve to a (differentiable) function u at a point $x_0 \in R^N$: we will use the notation $I_u(x_0) = \{x \in R^N | u(x) = u(x_0)\}$ to denote this indifference curve, which is generally a curved manifold of R^N, i.e. a "surface" in R^N that is one dimension less than N the dimension of the full space R^N. Thus when $N = 2$, an indifference curve is a *curved line* in R^2 and when $N = 3$ we get an *indifference surface* which is curved surface in R^3 and so forth (note a surface in R^3 is "locally" two dimensional, i.e. it is two dimensional instead of three-dimensional, i.e. not a "solid").

Now by Taylor's Theorem, if $u(x)$ is differentiable at a point x_0 we can form a *linear approximation* to the function $u(x)$ at a point x_0 via the *linear* function $l(x)$ given by

$$l(x) = u(x_0) + \langle \nabla(x_0), (x - x_0)\rangle. \tag{2}$$

Notice that this linear approximation $l(x)$ satisfies a) $l(x_0) = u(x_0)$ (i.e. the linear approximation to $u(x)$ *touches* u at the point x_0, and b) $\nabla l(x) = \nabla u(x_0)$, i.e. the *slope* of the linear function $l(x)$

equals the slope of the nonlinear function $u(x)$ at the point x_0 (which is just the gradient, $\nabla u(x_0)$). So now we can define the *tangent plane* to the function $u(x)$ at the point x_0 as the indifference curve for the function $l(x)$ at the point x_0

$$I_l(x_0) = \{x \in R^N | l(x) = l(x_0) = u(x_0)\} \tag{3}$$

Clearly the indifference curve of a linear function will just be a linear space, i.e. a *hyperplane*, rather than a curved surface or manifold which is what an indifference curve of a *nonlinear function* typically is.

So now it just boils down to showing that any point x in the tangent hyperplane is orthogonal to the gradient of $u(x)$ at x_0. But by the definition of the tangent hyperplane, $x \in I_l(x_0)$ if and only if we have

$$l(x) = u(x_0) + \langle \nabla u(x_0), (x - x_0) \rangle = l(x_0) = u(x_0). \tag{4}$$

But subtracting $u(x_0)$ on each side of the equation above, we see this is equivalent to

$$\langle \nabla u(x_0), (x - x_0) \rangle = 0. \tag{5}$$

This is *almost* the result we want, since equation (5) tells us that the point $(x - x_0)$ is orthogonal (perpendicular) to $\nabla u(x_0)$. But notice that the indifference curve $I_l(x_0)$ does not generally go through the origin, i.e. it is generally not the case that $0 \in I_l(x_0)$ (make sure you understand why this is the case). However we can define a *parallel shift* of this indifference curve by subtracting x_0 from every point in $I_l(x_0)$. That is, we can define the tangent hyperplane as a *parallel shift* of the tangent plane $I_l(x_0)$, so we define this tangent hyperplane by $T_l(x_0)$ as follows

$$T_l(x_0) = \{x \in R^N | x = (y - x_0) \text{ for some } y \in I_l(x_0)\}. \tag{6}$$

Notice that since $x_0 \in I_l(x_0)$ (x_0 is just the point of tangency), it follows that $x = (x_0 - x_0) = 0 \in T_l(x_0)$, so the tangent hyperplane includes the point $0 \in R^N$, and we say that the *tangent hyperplane passes through the origin.* Since any point $y \in I_l(x_0)$ satisfies the condition $\langle \nabla u(x_0), (y - x_0) \rangle = 0$, it follows that the point $x = (y - x_0) \in T_l(x_0)$ satisfies $\langle \nabla u(x_0), x \rangle = 0$. That is, we have shown that *any point in the tangent hyperplane to the function $u(x)$ at a point $x_0 \in R^N$ is orthogonal to the gradient* $\nabla u(x_0)$. It should also be clear why $T_l(x_0)$ is a hyperplane. As we discussed above, a hyperplane is any set of points in R^N that satisfy $\langle a, x \rangle = 0$ for some $a \in R^N$. But we have just shown that $T_l(x_0)$ is the set of all $x \in R^N$ such that $\langle \nabla u(x_0), x \rangle = 0$ so letting $a = \nabla u(x_0)$ it follows that $T_l(x_0)$ is indeed a hyperplane.

2. **Lagrangian saddlepoint solution for constrained optimization problems** Consider the following *constrained optimization problem*

$$\max_x u(x) \quad \text{subject to} \quad g(x) \geq 0, \quad \text{and} \quad x \geq 0 \tag{7}$$

where $u : R^n \to R$ is a continuous function and $g : R^n \to R^m$ are m *constraint functions* which are also continuous functions of x. Define the *Lagrangian* $\mathcal{L}(x, \lambda) : R^(n + m) \to R$ by

$$\mathcal{L}(x, \lambda) = u(x) + \lambda' g(x) \tag{8}$$

where

$$\lambda' g(x) = \langle \lambda, g(x) \rangle = \sum_{j=1}^{m} \lambda_j g_j(x). \tag{9}$$

Definition (x^*,λ^*) is a *saddlepoint* of $\mathcal{L}$ if and only if

$$\begin{array}{rcll} \mathcal{L}(x^*,\lambda^*) & \geq & \mathcal{L}(x,\lambda^*) & \forall x \geq 0 \\ \mathcal{L}(x^*,\lambda^*) & \leq & \mathcal{L}(x^*,\lambda) & \forall \lambda \geq 0 \end{array}$$

Theorem If (x^*,λ^*) is a saddlepoint of $\mathcal{L}$ then x^* solves the constrained optimization problem (7). Prove this theorem. **Hint:** use the method of *proof by contradiction.*

> **Proof** I proved this in class, perhaps not in every detail but I used a proof by contradiction to show that if (x^*,λ^*) is a saddlepoint of $\mathcal{L}(x,\lambda)$ then x^* must solve the constrained optimization problem (7). The more difficult thing to prove is the *converse* to this Theorem, i.e. if x^* solves the the constrained optimization problem (7), then there exists a $\lambda^* \geq 0$ such that (x^*,λ^*) is a saddlepoint to the Lagrangian $\mathcal{L}(x,\lambda)$. I *did not* ask you to prove this converse result and it is not true in general without *more assumptions.* It can be proven under the stronger assumptions that $\{x|g(x) \geq 0\}$ is a *convex set* and $u(x)$ is *quasiconcave.* Then we can appeal to a theorem called the *separating hyperplane theorem* to prove the existence of a Lagrange multiplier vector λ^* such that (x^*,λ^*) is a saddlepoint of $\mathcal{L}(x,\lambda)$. However this is beyond the level of this class and I emphasize it is *not* something I asked for or expected any of you to prove.

3. **Prove that if x^* is an *interior solution* that maximizes the consumer's problem below, the indifference curve at x^* is tangent to the budget line.**

$$\max_{x\geq 0} u(x) \quad \text{subject to} \quad \langle p,x\rangle \leq y \tag{10}$$

where $u: R^n \to R$ is a continuously differentiable utility function and $p \in R^n$ are positive prices of the n goods entering the consumer's utility function.

> **Answer** We need a slightly stronger assumption for this result to hold, namely we need to assume that *more is always better* which is mathematically equivalent to $\nabla u(x) > 0$ for any $x \geq 0$ (where the vector inequality $\nabla u(x) > 0$ means that each component of $\nabla u(x)$ is strictly greater than zero, so that $\frac{\partial}{\partial x_i}u(x) > 0$ for $i = 1,\dots,n$). Suppose the conditions to the converse to theorem above holds, i.e. there exists a $\lambda^* \geq 0$ such that (x^*,λ^*) is a saddlepoint to the Lagrangian $\mathcal{L}(x,\lambda)$. Since $u(x)$ is differentiable and (x^*,λ^*) is a saddlepoint, x^* must maximize $\mathcal{L}(x,\lambda^*)$ in x. Since x^* is interior — i.e. $x^* > 0$ — it follows that the the gradient of $\mathcal{L}(x,\lambda^*)$ with respect to x must be identically 0 at $x = x^*$, i.e. we must have
>
> $$\frac{\partial}{\partial x}\mathcal{L}(x,\lambda^*) = \nabla u(x^*) - \lambda^* p = 0, \tag{11}$$
>
> where 0 is interpreted as the zero vector in R^N. Since $\nabla u(x^*) > 0$ (by the "more is always better" assumption) and $p > 0$, it follows that $\lambda^* > 0$. Now we already showed in problem 1 above that the tangent hyperplane to the indifference curve of u at x^*, $I_u(x^*)$, is perpendicular (orthogonal) to $\nabla u(x^*)$ (i.e. it is the set of all $x \in R^N$ satisfying $\langle \nabla u(x^*),x\rangle = 0$). Now we show that the shifted budget line (the budget hyperplane) is also orthogonal to $\nabla u(x^*)$, but if both the tangent hyperplane to the indifference curve and the shifted budget line are both orthogonal to $\nabla u(x^*)$, then these must be the *same hyperplane* and thus *parallel* and thus, it follows that the budget line (or plane) must be tangent to the indifference curve $I_u(x^*)$. That is, we can define a shifted version of the budget line by the linear function $l(x)$ given by
>
> $$l(x) = u(x^*) + \langle \lambda^* p, (x - x^*)\rangle \tag{12}$$

Notice that a) $l(x)$ touches the indifference curve $I_u(x^*)$ since we have $l(x^*) = u(x^*)$, and b) the slope of $l(x)$ is the same as the slope of $u(x)$ at x^*. To see this latter point, note that the slope of $l(x)$ is just its gradient, which is $\nabla l(x) = \lambda^* \nabla p$. However by equation (11) we have $\lambda^* p = \nabla u(x^*)$. So it follows that the slope (gradient) of $l(x)$, which is just the slope of the (shifted) budget line, equals the slope of the indifference curve at $u(x^*)$ which is $\nabla u(x^*)$. In other words, the budget line is tangent to the indifference curve at the optimal bundle x^*.

4. **Firm Profit Maximization Problem** Consider a firm whose production function has 2 outputs, y_1 and y_2 and 2 inputs, x_1 and x_2. Suppose that its production function is given by

$$\left[y_1^2 + 4y_2^2\right]^{1/2} = \left[{x_1}^2 + {x_2}^2\right]^{1/2} \tag{13}$$

a. Does this production function have increasing, decreasing, or constant returns to scale? (Hint: if you double both inputs x_1 and x_2 can you double, more than double, or less than double both of the outputs y_1 and y_2?)

Answer: We can write a general production function as $F(y,x) \le 0$. The general definition of constant returns to scale is that if (y,x) is feasible to produce, i.e. if $F(y,x) \le 0$, then for any positive scalar $\lambda \ge 0$ we have that it is also feasible to produce $(\lambda y, \lambda x)$, i.e. we need to check that $F(\lambda y, \lambda x) \le 0$. For this problem it is easy to see that there are constant returns to scale since

$$F(y,x) = \left[y_1^2 + 4y_2^2\right]^{1/2} - \left[{x_1}^2 + {x_2}^2\right]^{1/2} \tag{14}$$

and it is easy to check that for any $\lambda \ge 0$ we have $F(\lambda y, \lambda x) \le 0$.

b. Suppose for a moment that we fix input levels so that $x_1 = x_2 = 5$. Plot the *output possibility frontier*, i.e. plot (in (y_2, y_1) space) the set of feasible combinations of y_1 and y_2 that can be produced using inputs $x_1 = x_2 = 5$.

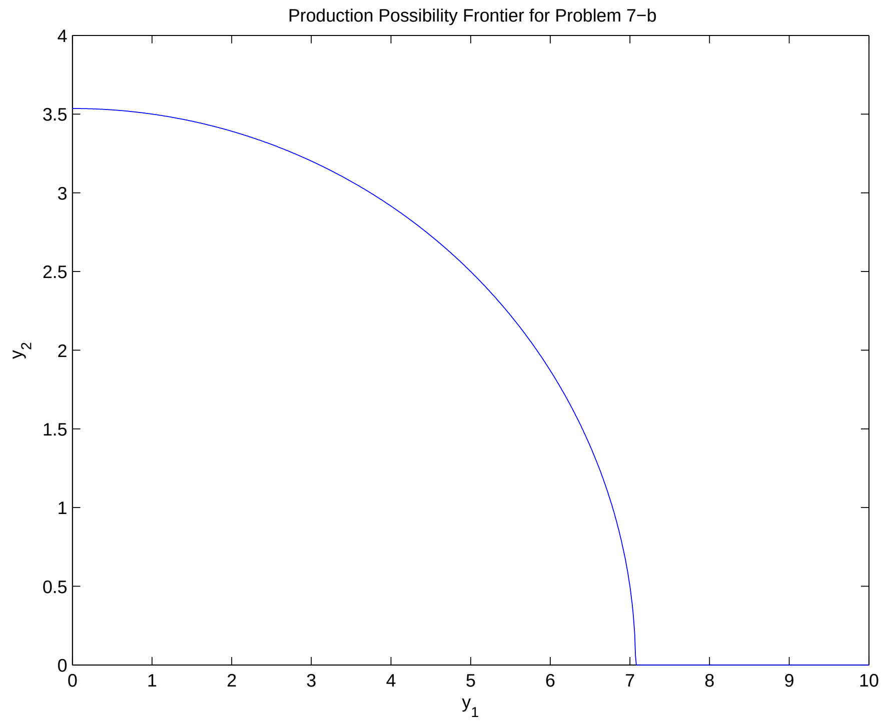
Answer: If $x_1 = x_2 = 5$ then the production function constraint tells us that the set of feasible outputs (y_1, y_2) that can be produced is

$$P = \left\{ (y_1, y_2) | \left[y_1^2 + 4y_2^2\right]^{1/2} \le \sqrt{50},\, y_1 \ge 0,\, y_2 \ge 0 \right\}. \tag{15}$$

This region is an *ellipse* and the production possibility frontier is the graph of the equation $y_1^2 + 4y_s^2 = \sqrt{50}$ and is plotted in figure 1 below. We can rewrite the equationi for the production possibility frontier as

$$y_2 = \sqrt{\frac{50 - y_1^2}{4}}. \tag{16}$$

Figure 1: Output Possibility Frontier for inputs $x_1 = x_2 = 5$



c. Continuing the previous question, if the output prices are $p_1 = 6$ for y_1 and $p_2 = 16$ for y_2, and if we assume that the inputs x_1 and x_2 are fixed at 5, what combination of outputs (y_1^*, y_2^*) maximize the firm's revenue? If we were to increase x_1 by a small amount, on the margin, by how much would the revenues of the firm increase (i.e. how much does revenue increase for an increase of amount ε, some small positive number, in input x_1)?

Answer: We want to maximize revenues subject to the fixed input constraints that $x_1 = x_2 = 5$. The Lagrangian for this problem is

$$\mathcal{L}(y_1, y_2, \lambda) = p_1 y_1 + p_2 y_2 + \lambda \left(\sqrt{5^2 + 5^2})^{1/2} - \sqrt{y_1^2 + 4y_2^2} \right) \tag{17}$$

I am going to leave it to you to write the first order conditions and solve them to get the optimal outputs (y_1^*, y_2^*) which are given by

$$\begin{aligned} y_1^* &= \frac{\sqrt{50}}{\sqrt{1 + \frac{p_2^2}{4p_1^2}}} \\ y_2^* &= \frac{\sqrt{50}}{\sqrt{4 + \frac{16p_1^2}{p_2^2}}} \end{aligned} \tag{18}$$

Plugging in the prices $p_1 = 6$ and $p_2 = 16$ into these formulas, we get

$$\begin{aligned} y_1^* &= 4.2426 \\ y_2^* &= 2.8284 \end{aligned} \tag{19}$$

and the maximized value of revenue is

$$R^* = p_1 y_1^* + p_2 y_2^* = 6 \times 4.2426 + 16 \times 2.8284 = 10\sqrt{50} = 70.71. \quad (20)$$

The final part of this question is to compute by how much revenue increases if input $x_1 = 5$ increases by a small amount ε. This can be computed as

$$\frac{\partial R^*}{\partial x_1}\varepsilon = \left[p_1 \frac{\partial y_1^*}{\partial p_1} + p_2 \frac{\partial y_2^*}{\partial p_2}\right]\varepsilon. \quad (21)$$

Using equation (18) we can compute the derivatives of y_1^* and y_2^* with respect to x_1 to get:

$$\begin{aligned} \frac{\partial R^*}{\partial x_1}\varepsilon &= \left[p_1 \frac{\partial y_1^*}{\partial p_1} + p_2 \frac{\partial y_2^*}{\partial p_2}\right]\varepsilon \\ &= \frac{1}{10}R^*\varepsilon \\ &= \sqrt{50}\varepsilon \\ &= 7.071\varepsilon, \end{aligned} \quad (22)$$

since $\partial y_1^*/\partial x_1 = \partial y_2^*/\partial x_1 = y_1^* x_1/50 = y_2^* x_1/50 = 1/10$ when $x_1 = 5$. We can also use the Lagrange multiplier from the Lagrangian to get this. From the first order condition for y_1^* we get

$$\lambda^* = \sqrt{50}\frac{p_1}{y_1^*} = \sqrt{50}\frac{6}{4.2426} = 10.00. \quad (23)$$

However λ^* measures the effect of relaxing the constraint, $K\sqrt{50}$, but we are interested in measuring the effect of increasing x_1 by ε. Using the chain rule,

$$\frac{\partial\sqrt{x_1^2 + x_2^2}}{\partial x_1} = \frac{x_1}{\sqrt{x_1^2 + x_2^2}} = \frac{5}{\sqrt{50}} \quad (24)$$

when $x_1 = x_2 = 5$. Thus, the effect of an increase of ε in x_1 on revenues using the Lagrange multiplier is

$$\frac{\partial R^*}{\partial x_1} = \lambda^* \frac{\partial\sqrt{x_1^2 + x_2^2}}{\partial x_1} = \frac{\lambda^* 5}{\sqrt{50}} = \sqrt{50} = 7.071. \quad (25)$$

So we get the same answer regardless of which route we take to computing it. That's reassurring!

d. Now consider what the optimal level of inputs should be in order to produce the (y_1^*, y_2^*) combination that you computed in part c. If the price of the inputs are $w_1 = 4$ for x_1 and $w_2 = 6$ for x_2, what is the cost-minimizing level of inputs that can produce (y_1^*, y_2^*)? (Hint: recall that when $x_1 = x_2 = 5$ we have $[{x_1}^2 + {x_2}^2]^{1/2} = \sqrt{50}$. So you need to minimize total costs $w_1x_1 + w_2x_2$ subject to the constraint that $[{x_1}^2 + {x_2}^2]^{1/2} = \sqrt{50}$). If we needed to increase output y_1 by a small amount, say by .1, approximately how much would it cost the firm to do this?

Answer: This answer is the same as the answer to problem 6: given that the input isoquants are concave rather than convex to the origin, the optimal policy is to use the cheaper of the two inputs. Thus, let w denote the cheaper of the two input prices and let $x =$ denote the quantity of the cheaper input that was used to produce the goods. In this case good 1 is cheaper since its price is $p_1 = 4$ and good 2's price is $p_2 = 6$. So the level of x_1 used is $x_1 = \sqrt{50} = 7.071$. To compute the additional cost of increasing y_1 by a small amount we need to consider the different ways we could increase y_1. One way would be to decrease the production of y_2 to increase the production of y_1, leaving the input requirement x unchanged. The other way is to assume that we increase the level of input 1, x. But this results in joint production of not only y_1 but also y_2. To measure the incremental cost properly, we have to deduct the increased revenues resulting from the use of the addition inputs. Computing the incremental cost the first way (i.e. holding x constant and treating the cost as the opportunity cost of lost sales of good 2). Totally differentiating the production function constraint we bet

$$\frac{y_1 dy_1}{\sqrt{y_1^2 + 4y_2^2}} + \frac{4y_2 dy_2}{\sqrt{y_1^2 + y_2^2}} = 0. \tag{26}$$

and solving for dy_2/dy_1 we get

$$\frac{dy_2}{dy_1} = -\frac{y_1}{4y_2} \tag{27}$$

so the opportunity cost in terms of lost output of y_2 from increasing y_1 by a small amount ε is $\varepsilon/4$ of the ratio y_1/y_2. If the firm is maximizing revenues, then we saw above that the optimal y_1^* and y_2^* are produced in the ratio of $4p_1/p_2$. With $p_1 = 6$ and $p_2 = 16$, then we have

$$\frac{dy_2}{dy_1} = -\frac{y_1^*}{4y_2^*} = -\frac{p_1}{p_2} = -\frac{6}{16}. \tag{28}$$

Since the price of good 2 is $p_2 = 16$, the cost to the firm of increasing y_1 by ε units is

$$p_2 \frac{dy_2^*}{dy_1^*}\varepsilon = -6\varepsilon. \tag{29}$$

You should show that the incremental cost will be the same if the firm decides to produce more of y_1 by increasing its input level x. However be careful to deduct the extra revenue from sales of increased amount of output of good 2 from the cost of the extra inputs!

e. Now step back and look at the firm overall. Is the production plan $(y_1^*, y_2^*, x_1^*, x_2^*)$ that you computed in parts c and d above a profit maximizing production plan for this firm? Why or why not?

Answer: Since the firm's production function has constant returns to scale, the profit maximizing scale of operations is not well-defined: if the firm can make positive profits at a given scale of operations, then it could increase profits without bounds by scaling up its levels of inputs and outputs simultaneously. If the prices are such that the firm earns zero profits at one scale of operation, then it is not hard to show that it will earn zero profits at any other scale of operations, so in either case, the scale of the firm is not well defined. So the only thing we can do is to determine the optimal combination of outputs for any arbitrarily fixed scale of operations. We can fix the scale by setting the cheaper of the two inputs, x to a given level such as $x = \sqrt{50}$. With this (arbitrary normalization), the optimal outputs computed in part a. are optimal, and thus part of a profit maximizing plan, provided profits are positive. Profit is

$$\Pi^* = p_1 y_1^* + p_2 y_2^* - wx = R^* - wx = 10\sqrt{50} - w\sqrt{10}. \tag{30}$$

Thus as long as $w \leq \sqrt{50} = 7.071$ the firm makes positive profits, and if it could, it would want to expand its scale of operations without bound to drive its profits to ∞. If $w = \sqrt{50}$ the firm makes

0 profits regardless of its scale of operations. If $w < \sqrt{50}$ then the firm would make a loss at any scale of operations, so its best course of action is to shut down.

Another way to solve this problem is to write down the Lagrangian to the firm's full profit maximization problem. The Lagrangian is

$$\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) = p_1y_1 + p_2y_2 - w_1x_1 - w_2x_2 + \lambda\left(\sqrt{w_1^2+w_2^2} - \sqrt{y_1^2+4y_2^2}\right) \tag{31}$$

Actually the algebra becomes a *lot easier* if we square both sides of the production function constraint and write it as

$$y_1^2 + 4y_2^2 = {x_1}^2 + {x_2}^2 \tag{32}$$

This is equivalent to the original production function constraint and the Lagrangian for the profit maximization problem with this simpler but equivalent version of the constraint is

$$\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) = p_1y_1 + p_2y_2 - w_1x_1 - w_2x_2 + \lambda\left(x_1^2 + x_2^2 - y_1^2 - 4y_2^2\right). \tag{33}$$

Now, recall that we want to maximize the Lagrangian over the variables (y_1,y_2,x_1,x_2) but to *minimize* it over the λ variable. The first order conditions for the maximization of $\mathcal{L}(y_1,y_2,x_1,x_2,\lambda)$ with respect to (y_1,y_2,x_1,x_2) are

$$\begin{aligned}
\frac{\partial}{\partial y_1}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= p_1 - 2\lambda y_1 \le 0\\
\frac{\partial}{\partial y_2}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= p_2 - 8\lambda y_2 \le 0\\
\frac{\partial}{\partial x_1}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= w_1 + 2\lambda x_1 \le 0\\
\frac{\partial}{\partial x_2}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= w_2 + 2\lambda x_2 \le 0.
\end{aligned} \tag{34}$$

I have written the first order conditions as *inequalities* in equation (34) to account for the possibility of *corner solutions.* We have already been alerted to this possibility in part b above, where we plotted the isoquants in (x_1,x_2) space and showed they were *concave* and thus the input cost-minimizing bundle would be to let either $x_1 = 0$ or $x_2 = 0$ depending on which of the two inputs is more expensive. When $w_1 \neq w_2$ we can see from the first order conditions for the Lagrangian in (34) above that it is impossible to have an *interior solution* in (x_1,x_2) (i.e. where $x_1 > 0$ and $x_2 > 0$ *simultaneously*). To see this, if there was an interior solution, then we would be able to take a second derivative of the Lagrangian to check the *second order conditions.* We would find the following:

$$\begin{aligned}
\frac{\partial^2}{\partial^2 y_1}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= -2\lambda \le 0\\
\frac{\partial}{\partial y_2}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= -8\lambda \le 0\\
\frac{\partial}{\partial x_1}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= 2\lambda \ge 0\\
\frac{\partial}{\partial x_2}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= 2\lambda \ge 0.
\end{aligned} \tag{35}$$

The second order conditions tell us that while there can be an interior solution for (y_1,y_2) (since the second derivative of the Lagrangian with respect to y_1 and y_2 are both negative if $\lambda > 0$), there cannot be an interior solution for (x_1,x_2) when $\lambda > 0$ since then the second order conditions for x_1 and x_2 are *positive* indicating that these would constitute a *local minimum* of the Lagrangian, but we are looking for x_1 and x_2 that *maximize* the Lagrangian. So we conclude that if $\lambda > 0$ then the only possible solution is for $x_1 = x_2 = 0$ and the first does not produce anything and earns zero profit. However it is also possible that $\lambda = 0$. In this case the second order condition in equation (35) above would be zero, which is not necessarily any contradiction, but the first order conditions in equation (34) will no longer hold.

So this is a very tricky problem where the Lagrangian approach is not that useful. To see what the general solution is we need to take another tack, which is to use the insight from part b above that the firm will set $x_2 = 0$ if $w_2 > w_1$ and set $x_1 = 0$ if $w_1 > w_2$. Let x denote the amount of the cheaper input that the firm uses and let $w = \min[w_1,w_2]$. For a fixed level of inputs we can solve a *conditional profit maximization problem* namely

$$\max_{y_1,y_2} p_1y_1 + p_2y_2 - wx \quad \text{subject to:} \quad x^2 = y_1^2 + 4y_2^2 \tag{36}$$

Call the solution to this problem $\Pi(p_1, p2, w|x)$ the *conditional profit function* since it is conditional on the firm restricting its inputs to the level x. What we want to do now is solve this conditional profit maximization problem and find an expression for $\Pi(p_1, p_2, w|x)$. Then we can do a *second stage optimization* to find the optimal level of the input x to maximize profits. The Lagrangian for the first stage conditional profit maximization problem is

$$\mathcal{L}(y_1,y_2,\lambda) = p_1y_1 + p_2y_2 - wx + \lambda\left(x^2 - y_1^2 - 4y_2^2\right). \tag{37}$$

The first order conditions are

$$\begin{aligned}
\frac{\partial}{\partial y_1}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= p_1 - 2\lambda y_1 = 0 \\
\frac{\partial}{\partial y_2}\mathcal{L}(y_1,y_2,x_1,x_2,\lambda) &= p_2 - 8\lambda y_2 = 0.
\end{aligned} \tag{38}$$

Now we can solve these equations to get $y_1 = p_1/2\lambda$ and $y_2 = p_2/8\lambda$. We can substitute these into the production function constraint $x^2 = y_1^2 + 4y_2^2$ to solve for λ

$$x^2 = \left[\left(\frac{p_1}{2\lambda}\right)^2 + \left(\frac{p_2}{8\lambda}\right)^2\right] \tag{39}$$

Solving for λ we get

$$\lambda = \frac{\sqrt{p_1^2 + p_2^2/4}}{2x}. \tag{40}$$

Substituting this equation for λ into the equations for y_1 and y_2 above we get

$$\begin{aligned}
y_1 &= \frac{xp_1}{\sqrt{p_1^2 + p_2^2/4}} \\
y_2 &= \frac{xp_2}{4\sqrt{p_1^2 + p_2^2/4}}.
\end{aligned} \tag{41}$$

Now, substituting these formulas for the optimal outputs results in the following formula for the conditional profit function $\Pi(p_1, p_2, w|x)$

$$\Pi(p_1, p_2, w|x) = x\left[\sqrt{p_2^2 + p_2^2/4} - w\right]. \tag{42}$$

Now, if the term in brackets in equation (42) is strictly positive, the firm would want to increase production without bound, and drive the input level x to infinity. But infinite inputs and infinite profits is not a legitimate solution. If the term in brackets is zero, then the firm makes zero profits regardless of the scale of production x and it does not care what value x would be, and technically there are a *continuum* of profit maximizing solutions in this case. If the term in brackets in equation (42) is negative, then the firm wants to set $x = 0$ and so it does not produce anything and makes zero profits.

f. **Super bonus question:** If you answered in part e that the production plan $(y_1^*, y_2^*, x_1^*, x_2^*)$ computed in parts c and d is not a profit maximizing plan, then find the profit maximizing production plan.

Answer: I described the profit maxmizing plan and showed that the optimal revenue-maximizing output combination from part c. is profit maximizing provided w is below $\sqrt{50}$. If $w = \sqrt{50}$ then the firm gets zero profits as any scale of production also yields 0 profits. However if w is strictly less than $\sqrt{50}$ then as I showed above, the firm would want to expand its scale without bound, so then the answer in part c is not optimal. If $w < \sqrt{50}$ then the firm would make 0 profits at any positive scale of production, so its best course of action is to shut down.

5. **Bertrand Duopoly Problem** Consider those regions in the Washington DC area where households have a choice between two cable tv/internet providers: Comcast and Starpower. Assume that these companies do not engage in price discrimination, but rather provide cable/internet using a simple single per month pricing scheme. Assume also that there are no switching or hookup costs, so that customers can switch from Starpower to Comcast or vice versa (or to not have cable) at zero cost. We now consider the pricing problem faced by these two competing customers, treating their services as imperfect substitutes in the minds of the consumers in the Washington DC area. Thus, a household in this area has the following television "mode" choices:

1. No pay TV (i.e. watch broadcast TV, or don't watch TV or use broadband)
2. cable TV/broadband (via Comcast)
3. cable TV/broadband (via Starpower)

Of course, it is possible for some households to subscribe to both Starpower and Comcast simultaneously, but I assume that this is too expensive relative to the incremental value of having both hooked up, so that virtually no households would subscribe to both at the same time. Thus, I have limited households to the 3 possible choices given above, which I assume are mutually exclusive and exhaustive (having ruleed out the possibility of subcribing to both Comcast and Starpower).

Assume that Starpower and Comcast choose their prices independently and without any collusion as part of a Nash equilibrium in which each tries to maximize its profits, treating the price of its opponent as given. Initially I ignore the presence of explicit or implicit regulatory constraints. I assume that in the DC area where these two companies provide overlapping coverage there are N households. Let $P_c(p_c, p_s)$ denote the fraction of these N households who choose Comcast, and $P_s(p_c, p_s)$ be the fraction who choose Starpower. The remaining fraction, $1-P_c(p_c,p_s)-P_s(p_c,p_s)$ either watch broadcast TV (which has a price of \$0 per month), or do not watch TV or need broadband internet at all (god forbid!). It is convenient to start with a simple logit representation for the market shares for Comcast and Starpower:

$$\begin{aligned} P_c(p_c,p_s) &= \frac{\exp\{a_c+b_cp_c\}}{1+\exp\{a_c+b_cp_c\}+\exp\{a_s+b_sp_s\}} \\ P_s(p_s,p_s) &= \frac{\exp\{a_s+b_sp_s\}}{1+\exp\{a_c+b_cp_c\}+\exp\{a_s+b_sp_s\}} \end{aligned} \tag{43}$$

A more advanced approach would derive these market shares from a household level demand study, using micro data to estimate the consumer choices and accounting for other demographic variables, including household income y, and the characteristics of the "outside alternative", i.e. the characteristics of free to air TV. I assume these market shares are "reduced forms" consistent with the results of a micro level study. This initial "reduced form" approach requires specification of 7 pieces of information in order to predict the prices, profits, and market shares for Comcast and Starpower:

1. the number of households N in the "overlap region" served by both Comcast and Starpower,
2. the 4 market share coefficients (a_o,b_o,a_f,b_f)
3. the 2 marginal cost parameters (k_c,k_s)

Given suggested values for these 7 parameters, your job is to compute the Bertrand Nash equilibruium outcome, i.e. the prices that Comcast and Starpower will charge, their profits, and their equilibrium market shares.

Let k_c and k_s denote the marginal costs (i.e. costs which depend on the number of their subscribers) of providing their cable service. Then the Nash equilibrium, profit maximization conditions determining the prices (p_c^*,p_s^*) (where the $*$ superscripts denote their Nash equilibrium values) are given by

$$\begin{aligned} p_c^* &= \underset{p_c}{argmax}(p_c-k_c)NP_c(p_c,p_s^*) \\ p_s^* &= \underset{p_s}{argmax}(p_s-k_s)NP_s(p_c^*,p_s) \end{aligned} \tag{44}$$

Note that I have treated the cost of the programming content that Comcast and Starpower purchase as fixed costs, F_c and F_s that do not depend on the number of customers and thus do not enter into the determination of the the equilibrium prices (p_o^*,p_f^*). This would change if Comcast and Starpower paid per subscriber royalty fees to HBO, ESPN, and the other providers of their programming content. These fees would then be embodied in the marginal cost parameters (k_c,k_s).

Figure 1 shows an illustrative Bertrand-Nash equilibrium calculated for a particular choice of the 5 parameters given above. Notice that the number of households N is just a multiplicative constant in the profit functions for Comcast and Starpower and thus, in actuality, the equilibrium is fully determined by the 6 parameters $(a_c,b_c,a_s,b_s,k_c,k_s)$.

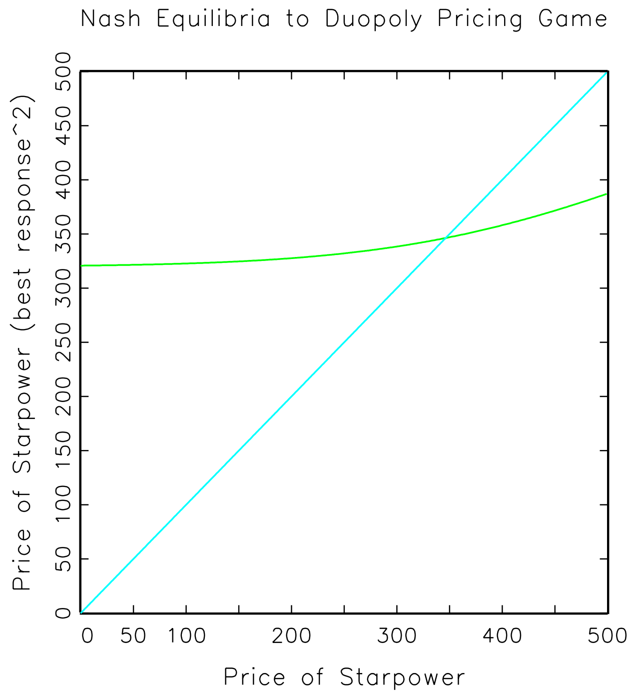
Your job is to try to calculate the equilibrium, writing the necessary programs to calculate the equilibrium in your favorite programming language. Once you calculate the equilibrium, prepare a plot of the equilibrium as done above and determine whether or not the equilibrium appears to be unique (inthe diagram above, it is clear that there is a unique "stable" equilibrium).

Also, I want you to compare the Bertrand-Nash duopoly outcome with the two possible monopoly outcomes:

1. Comcast has a monopoly in the DC area

2. Starpower has a monopoly in the DC area

Figure 1: Example of a Bertrand-Nash Equilibrium



Under the monopoly scenario, customers have only two options: 1) watch broadcast TV (or dont' watch), 2) subscribe to cable. If Comcast is the monopolist, assume that the share of the DC households it could obtain if it charged price p_c is given by

$$P_c(p_c) = \frac{\exp\{a_c + b_c p_c\}}{1 + \exp\{a_c + b_c p_c\}} \tag{45}$$

and if Starpower is the monopolist and charged price p_s it would get the following share of DC households

$$P_s(p_s) = \frac{\exp\{a_s + b_s p_s\}}{1 + \exp\{a_s + b_s p_s\}} \tag{46}$$

Thus, I assume that the same set of market share or "demand" coefficients (a_c, b_c) and (a_s, b_s) hold in the monopoly case as in the duopoly case.

Your job is to compute the monopoly and duopoly outcomes, and predict by how much cable prices would go up in the DC area if Comcast or Starpower gained monopoly control of this market.

To get you started I have provided a Gauss file `setup.gpr` that contains parameter values that you can use to compute the duopoly and monopoly outcomes, and a Gauss procedure, `br_c.g`, which computes the "best response function" for Comcast, i.e.

$$p_c = br_c(p_s) = \underset{p}{argmax}(p - k_c)NP_c(p, p_s) \tag{47}$$

which is Comcast's optimal price given that Starpower charges a price of p_s. With this hint you should be able to program the other pieces and compute the solution to this problem. You do not need to do your programming in Gauss: I have used Gauss only as an illustration to get you started.

Answer Using Matlab files that I have posted along with these answers on the Econ 425 web site, I obtain the following solutions. For the Bertrand duopoly, the Matlab function `equil.m` calculates Bertrand duopoly prices of $p_c = 173.89$ and $p_s = 346.30$. Comcast achieves a market share of 21.82% and Starpower has a market share of 71.12% and 0.76% of the people in this market do not use either Comcast or Starpower. The profits for Comcast are \$4.888 million and the profits for Starpower are \$24.630 million.

If Starpower is the only cable company, it would charge a price of \$632.72 and would earn \$53.27 million in profits. It would serve 84.2% of the households, but due to the high monopoly price, 15.8% of the households go without cable.

In the case where Comcast is the only cable company, it would charge a price of \$490.78 and earn profits of \$36.578 million. It would serve 74.5% of the market. Because of the perceived lower quality of Comcast's servicea, it cannot manage to charge as high of a monopoly price as Starpower can, and more consumers decide to go without cable when Comcast is the monopolist compared to when Starpower is the monopolist.

In the Bertrand case, the higher quality of service that Starpower provides it customers enables it to charge a significantly higher price and obtain a significantly larger market share than Comcast can obtain. However the competition between the two firms drives dow the prices to consumers by a significant amount, and the lower prices induces virtually all households to subscribe to cable.

6. **Intertemporal utility maximization with certain lifetimes.** Suppose a person has an additively separate, discounted utility function of the form

$$V(c_1,\ldots,c_T) = \sum_{t=1}^{T} \beta_s^t u(c_t) \tag{48}$$

where β_s is a subjective discount factor and $u(c_t)$ is an increasing utility function of consumption c_t in period t. Let the market discount factor is $\beta_m = 1/(1+r)$ where r is the market interest rate.

a. If $\beta_s = \beta_m$ show that the optimal consumption plan in a market where there are no borrowing constraints (i.e. the consumer has unlimited ability to borrow and lend subject to an intertemporal budget constraint) is to have a constant consumption stream over time, i.e. $c_1 = c_2 = \cdots = c_t = c_{t+1} = \cdots = c_T$.

b. If $\beta_s < \beta_m$ will the optimal consumption stream be flat, increasing over time, or decreasing over time, or can't you tell from the information given?

c. How does your answer to part b change if I tell you that the utility function $u(c)$ is convex in c?

Answers: The answer to this question is in my lecture notes. See pages 23 onward in the lecture notes on intertemporal choice. For part c, note that if the utility function is convex, then $u''(c) > 0$ and the answers to parts b is reversed, iv $\beta_s < \beta_m$, then optimal consumption will be *increasing* over time, the opposite of the case if utility is concave (diminishing marginal utility), in which case consumption would be *decreasing* over time.

7. **Consumption and Taxes** Suppose a consumer has a utility function $u(x_1, x_2) = \log(x_1) + \log(x_2)$ and an income of $y = 100$ and the prices of the two goods are $p_1 = 2$ and $p_2 = 3$.

a. In a world with no sales or income taxes, tell me how much of goods x_1 and x_2 this consumer will purchase.

answer Notice that the utility function is a monotonic transformation of a Cobb-Douglas utility function $l(x_1, x_2) = x_1^{1/2} x_2^{1/2}$, so demands are $x_1(p_1, p_2, y) = y/2p_1$ and $x_2(p_1, p_2, y) = y/2p_2$. With these, it is very easy to answer this question. $x_1 = 100/(2*2) = 25$ and $x_2 = 100/(2*3) = 16.66667$.

b. Now suppose there is a 10% a sales tax on good 1. That is, for every unit of good 1 the person buys, he/she has to pay a price of $p_1(1 + .1) = 2.2$, where the 10% of the price, or 20 cents, goes to the government as sales tax. How much of goods 1 and 2 does this person buy now?

answer With the tax in place, the price of good 1 increases to 2.2 so quantities demanded are $x_1 = 100/(2*2.2) = 22.727273$ and $x_2 = 100/(2*3) = 16.66667$. The total taxes the person pays are $.2x_1 = .2100/(2*2.2) = 4.54$.

c. Suppose instead there is a 5% income tax, so that the consumer must pay 5% of his/her income to the government. If there is no sales tax but a 5% income tax, how much of goods 1 and 2 will the consumer consume?

answer With a 5% income tax, the consumer has after-tax income equal to \$95 ($100(1-\tau)$ where $\tau = .05$). So the consumption of goods 1 and 2 is given by $x_1 = 95/(2*2) = 23.75$ and $x_2 = 95/(2*3) = 15.8333$.

d. Which would the consumer prefer, a 10% sales tax on good 1, or a 5% income tax? Explain your reasoning for full credit.

answer With the sales tax, the consumer consumes less of good 1 and more of good 2, and pays less in tax overall. With the income tax the consumer consumes more of good 1 but less of good 2 and pays more overall in tax (\$5.00 versus \$4.54). But the only way to see which alternative the consumer prefers is to plug the consumption bundles into his/her utility function and see which one give more utility. The utility under the sales tax is $\log(22.727273) + \log(16.66667) = 5.93699764$. The consumer's utility under the income tax is $\log(23.75) + \log(15.8333) = 5.9297$ so the consumer prefers the sales tax to the income tax.

e. How big would the sales tax on good 1 have to be for the government to get the same revenue as a 5% income tax? Which of the two taxes would the consumer prefer in this case, or is the consumer indifferent because the consumer has to pay a total tax of \$5 (5% of \$100) in either case?

answer Now we want to set the sale tax rate α so that we raise tax revenue of \$5, the same revenue that we collect under an income tax of 5%. The equation for the necessary tax rate is

$$5 = \alpha \frac{100}{2(2+\alpha)} \tag{49}$$

Solving this for α we get $\alpha = 2/9 = .22222$. Under this tax rate, consumption of good 1 falls to $x_1 = \frac{100}{2(2+\alpha)} = 22.5$ and the tax revenue collected is $22.5*2/9 = 5$. Now the person's utility under the sales tax is $\log(22.5) + \log(16.66667) = 5.926926$, so that now, the consumer slightly prefers to have the income tax over the sales tax.

8. **Supply and Demand Problem** The supply for corn is given by

$$S = 10 + 5p + .05R \tag{50}$$

where R is the amount of rainfall. The demand for corn is given by

$$D = 5Y^{.2}p^{-.5} \tag{51}$$

where Y is the per capita income.

a. What is the equation for the equilibrium price of corn, assuming this is a competitive market?

answer We find the price that sets supply equal to demand. Equilvalently, we seek a price p that sets *Excess demand* $E(p) = D(p) - S(p)$ to zero, where

$$E(p) = 5Y^{.2}p^{-.5} - (10 + 5p + .05R) \tag{52}$$

b. Solve for the equilibrium price and quantity in this market, using numerical methods (e.g. Newton's method) if necessary, or by any means possible to get numerical answers.

answer Suppose we set $Y = 1000$ and $R = 20$. Then the equation we want to solve is

$$E(p) = 5\left[1000^{.2}p^{-.5}\right] - 10 - 5p - 1 = 0. \tag{53}$$

I programmed this function in Matlab as the file `ed.m` which is posted on the Econ 625 website along with these answers. You can use Newton's method to solve this equation. I was a bit lazy and instead used the Matlab `fsolve` command to solve this equation, that is, I did `fsolve(@ed,2)` (so that my initial guess for a solutions was $p = 2$). `fsolve` returned the solution $p = 1.2964$ and checking, `ed(1.2964)`$= 1.4087e^{-10}$.

c. Derive a formula for dp/dR, i.e. the effect of an increase in rainfall on the price of corn.

Answer This is an exercise in the use of the implicit function theorem. See my lecture notes on this posted on the Econ 425 website. But the answer is

$$\frac{dp}{dR} = -\frac{.05}{2.5y^{.2}p^{-1.5} + 5} < 0. \tag{54}$$

d. Derive a formula for dp/dY, i.e. the effect of an increase in per capita income on the price of corn.

Answer Using the implicit function theorem again,

$$\frac{dp}{dy} = -\frac{\frac{\partial}{\partial y}E(p,y)}{\frac{\partial}{\partial p}E(p,y)} = \frac{y^{-.8}p^{-.5}}{2.5y^{.2}p^{-1.5} + 5} > 0. \tag{55}$$