

13. CONSTRAINED OPTIMIZATION: THE LAGRANGIAN METHOD

J.P. Bastos*

EPH 1100: Analytical Tools

Reading: Hoy et al., §13.1–13.2.

- **§13.1:** Constrained optimization; the Lagrangian function; first-order conditions; the Lagrange multiplier as a shadow price.
- **§13.2:** Second-order conditions for constrained optimization (the bordered Hessian).

1. Constrained Optimization in Economics

Our work so far has been focused on finding the optimal point in a function. However, most applications in economics involve optimizing *subject to a constraint*. A consumer must choose between goods, but she cannot spend more than her income; a firm must maximize profit, but it cannot produce more than its capacity allows.

Indeed, the two most fundamental problems in economics are constrained optimization problems. For instance, take the consumer’s problem:

Maximize utility $U(x_1, x_2)$ subject to the budget constraint $p_1x_1 + p_2x_2 = m$.

The consumer wants to reach the highest possible utility. But she cannot choose any bundle she likes; she is confined to the **budget line** $p_1x_1 + p_2x_2 = m$. Every bundle she can afford lies exactly at (or below) this line and bundles above it are out of reach — bundles below the line are affordable, but they will not *maximize* utility.¹

2. The Lagrangian

At first sight, the unconstrained methods we learned so far will be insufficient, insofar as we cannot simply set partial derivatives to zero and expect to land within the constraint. We must find the

*Assistant Professor of Economics, University of Austin (UATX). jbastos@uaustrin.org

¹In most models we assume the consumer spends all of their income, unless we are explicitly modeling savings, in which case one of the goods (x_1, x_2) can represent “future consumption.”

best point *on* the constraint line, not the best point *anywhere*. However, there is a simple and elegant method that allows us to solve constrained problems using the same tools we have already learned: the **Lagrangian method**.

The key insight is to transform the constrained problem into an unconstrained one by incorporating the constraint into the function we want to optimize, the **objective function**. We do this by introducing a new variable that connects the constraint to the objective function, the **Lagrange multiplier**, which measures how much the objective function would increase if we relaxed the constraint by one unit.

Definition. Given the problem

$$\max_{x,y} f(x,y) \quad \text{subject to} \quad g(x,y) = c$$

the **Lagrangian** is the function:

$$\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda[c - g(x,y)]$$

The new variable λ is the **Lagrange multiplier**.

Why does this work? The term $\lambda[c - g(x,y)]$ is zero whenever the constraint $g(x,y) = c$ is satisfied. So, at any feasible point, $\mathcal{L} = f$. The multiplier λ acts as a “penalty” for violating the constraint: it forces the unconstrained first-order conditions of $\mathcal{L}$ to coincide with the constrained optimum of f .

For the consumer problem, the Lagrangian is:

$$\mathcal{L}(x_1, x_2, \lambda) = U(x_1, x_2) + \lambda(m - p_1x_1 - p_2x_2)$$

For the producer’s cost-minimization problem, the Lagrangian is:

$$\mathcal{L}(L, K, \lambda) = wL + rK + \lambda[\bar{Q} - F(L, K)]$$

3. First-Order Conditions

To find the critical points of $\mathcal{L}$, differentiate with respect to each of the three variables x , y , and λ , and set all three derivatives to zero.

FOC for the Lagrangian. At the constrained optimum (x^*, y^*, λ^*) :

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= f_x - \lambda g_x = 0 \implies f_x = \lambda g_x \\ \frac{\partial \mathcal{L}}{\partial y} &= f_y - \lambda g_y = 0 \implies f_y = \lambda g_y \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= c - g(x, y) = 0 \implies g(x^*, y^*) = c\end{aligned}$$

Dividing the first condition by the second (when $g_y \neq 0$ and $\lambda \neq 0$):

$$\frac{f_x}{f_y} = \frac{g_x}{g_y}$$

For the consumer problem, this becomes:

$$\frac{U_1}{U_2} = \frac{p_1}{p_2}$$

The ratio U_1/U_2 is the **marginal rate of substitution (MRS)**: how many units of good 2 the consumer is willing to give up for one more unit of good 1. The ratio p_1/p_2 is the **price ratio**: the market rate at which goods can be exchanged. At the optimum, these must be equal.

Notice that you can also rearrange the terms to get:

$$\frac{U_1}{U_2} = \frac{p_1}{p_2} \iff \frac{U_1}{p_1} = \frac{U_2}{p_2}$$

The ratio U_1/p_1 is the **marginal utility per dollar** spent on good 1 (and likewise for good 2). At the optimum, the last dollar spent on either good buys the same extra utility. If good 1 delivered more utility per dollar than good 2, the consumer could raise her utility by shifting a dollar of spending from good 2 to good 1 — so an unequal ratio can never be an optimum.

The same method works for minimization problems too: minimizing $h(x, y)$ is the same problem as maximizing $-h(x, y)$, so the Lagrangian construction and the FOC above carry over unchanged. This is exactly what we need for a firm minimizing cost.

Analogously, for a firm minimizing cost subject to a production constraint, the FOC give:

$$\frac{F_L}{F_K} = \frac{w}{r}$$

The ratio F_L/F_K is the **marginal rate of technical substitution (MRTS)**: how many units of capital the firm is willing to give up for one more unit of labor. The ratio w/r is the **factor price ratio**: the market rates for *wages* and *rental rate of capital*, which also informs the rate at which inputs can be exchanged. At the optimum, these must be equal.

The same rearrangement applies here:

$$\frac{F_L}{F_K} = \frac{w}{r} \iff \frac{F_L}{w} = \frac{F_K}{r}$$

The ratio F_L/w is the **marginal product of labor per dollar** spent on labor (and likewise for capital). At the cost-minimizing input mix, every dollar spent buys the same amount of extra output, whether it is spent on labor or capital. If labor delivered more output per dollar than capital, the firm could produce the same output more cheaply by shifting spending toward labor.

4. The Lagrange Multiplier

The Lagrange multiplier λ has a concrete economic meaning that is worth understanding before doing any calculations. It measures the rate at which the optimized objective value increases when the constraint is relaxed by one unit. It is the **shadow price** of the constraint.

Definition. For the problem $\max_{x,y} f(x,y)$ subject to $g(x,y) = c$, the Lagrange multiplier λ^* is the **shadow price** of the constraint: it measures the rate at which the optimized value of the objective function increases when the constraint level c is relaxed by one unit.

$$\lambda^* \approx \frac{\Delta f^*}{\Delta c}$$

This is a general result; what λ^* *means* depends on what the objective and constraint represent in the problem at hand.

Applied to the consumer's problem, the constraint level is income m , so $\lambda^* = \frac{\Delta U^*}{\Delta m}$ is the **marginal utility of income**: the extra utility obtained from one additional dollar of budget.

Applied to the firm's cost-minimization problem, the constraint level is the output target $\bar{q}$, so $\lambda^* = \frac{\Delta C^*}{\Delta \bar{q}}$ is the **marginal cost of output**: the extra cost incurred from producing one additional unit of output.

Example. Consider a consumer with utility $U(x_1, x_2) = x_1^{0.5}x_2^{0.5}$, prices $p_1 = 2$ and $p_2 = 4$, and income $m = 100$. We will solve the problem, analyze and interpret λ^* , and then verify by re-solving with $m = 101$.

Step 1: Write the Lagrangian and the FOC.

$$\mathcal{L} = x_1^{0.5}x_2^{0.5} + \lambda(100 - 2x_1 - 4x_2)$$

$$\mathcal{L}_{x_1} = 0.5 x_1^{-0.5} x_2^{0.5} - 2\lambda = 0 \tag{1}$$

$$\mathcal{L}_{x_2} = 0.5 x_1^{0.5} x_2^{-0.5} - 4\lambda = 0 \tag{2}$$

$$\mathcal{L}_\lambda = 100 - 2x_1 - 4x_2 = 0 \tag{3}$$

Step 2: Take (1) and (2), and solve each one for λ and set the two expressions equal to each other, which eliminates λ .²

$$\text{From (1):} \quad \lambda = 0.25 x_1^{-0.5} x_2^{0.5}$$

$$\text{From (2):} \quad \lambda = 0.125 x_1^{0.5} x_2^{-0.5}$$

Setting these equal to each other gives the MRS condition:

$$0.25 x_1^{-0.5} x_2^{0.5} = 0.125 x_1^{0.5} x_2^{-0.5}$$

Now solve for x_1 in terms of x_2 : gather the x_2 terms on one side and the x_1 terms on the other, so the exponents cancel directly.

$$\frac{x_2^{0.5}}{x_2^{-0.5}} = \frac{0.125}{0.25} \cdot \frac{x_1^{0.5}}{x_1^{-0.5}} \implies x_2 = 0.5 x_1$$

Written as a ratio, this is exactly the MRS = price ratio condition from Section 3, now with numbers plugged in:

$$\underbrace{\frac{x_2}{x_1}}_{\text{MRS} = U_{x_1}/U_{x_2}} = \underbrace{0.5}_{\text{price ratio} = p_1/p_2} \implies x_1 = 2x_2$$

This says the consumer buys twice as much of good 1 as of good 2. Since good 1 costs half as much as good 2 ($p_1 = 2$ vs. $p_2 = 4$), and the utility function weighs the two goods equally (notice the exponents are the same), the consumer buys proportionally more of the cheaper one.

Step 3: Substitute $x_1 = 2x_2$ (or $x_2 = 0.5x_1$) into the budget constraint (3) and solve for x_1 and x_2 to find the quantity demanded of each good at these prices and this income.

$$2(2x_2) + 4x_2 = 100 \implies 8x_2 = 100 \implies x_2^* = 12.5, \quad x_1^* = 25$$

These are the consumer's demand for goods 1 and 2: at $p_1 = 2$, $p_2 = 4$, $m = 100$, she buys 25 units of good 1 and 12.5 units of good 2.

Step 4: Find λ^* .

Condition (3), $\partial \mathcal{L} / \partial \lambda = 0$, cannot help us here: it is just the budget constraint written out again, $100 - 2x_1 - 4x_2 = 0$, and λ does not even appear in it. We already used it in Step 3 to solve for x_1 and x_2 ; it has nothing left to say about λ .

To find λ , we need one of the *other* two conditions, (1) or (2), since those are the only ones

²There are many ways to solve this, but we typically do it this way to show the MRS condition explicitly, and to show both sides of the problem (in terms of x_1 and x_2), which is useful for interpretation and for deriving the demand functions later.

that actually contain λ . Take (1) and solve it for λ :

$$\lambda = \frac{U_{x_1}(x_1, x_2)}{p_1} = \frac{0.5 x_1^{-0.5} x_2^{0.5}}{2} = 0.25 x_1^{-0.5} x_2^{0.5}$$

Notice U_{x_1} , the marginal utility of good 1, is a function of *both* x_1 and x_2 : how much extra satisfaction one more unit of good 1 delivers can depend on how much of good 2 you already have. So λ is itself a function of the bundle (x_1, x_2) — it is not a fixed number until we say which bundle we mean.

Could we have used (2) instead? Yes. It gives $\lambda = U_{x_2}/p_2 = 0.125 x_1^{0.5} x_2^{-0.5}$, and at the optimum this must equal the expression above — that is exactly the MRS condition from Step 2. Either condition gives the same λ^* once we plug in the optimal bundle; we use (1) below and check the answer against (2).

Evaluate at the optimal bundle $(x_1^*, x_2^*) = (25, 12.5)$ from Step 3. Plug it into the expression for λ from (1):

$$\lambda^* = 0.25 (x_1^*)^{-0.5} (x_2^*)^{0.5} = 0.25 \times 25^{-0.5} \times 12.5^{0.5} = 0.25 \times 0.2 \times 3.536 \approx 0.177$$

As a check, plug the same bundle into the expression for λ from (2):

$$\lambda^* = 0.125 (x_1^*)^{0.5} (x_2^*)^{-0.5} = 0.125 \times 25^{0.5} \times 12.5^{-0.5} = 0.125 \times 5 \times 0.283 \approx 0.177$$

Same answer, as it must be.

$$\text{Optimal utility: } U^* = 25^{0.5} \times 12.5^{0.5} = \sqrt{312.5} \approx 17.678.$$

Step 5: Now assume m increases from 100 to 101 and then re-solve. The condition $x_1 = 2x_2$ is unchanged (it does not depend on m). The new budget constraint gives $x_2^* = 101/8 = 12.625$ and $x_1^* = 101/4 = 25.25$.

$$\text{New utility: } U^* = \sqrt{25.25 \times 12.625} = \sqrt{318.78125} \approx 17.854.$$

$$\text{Change in utility: } \Delta U \approx 17.854 - 17.678 = 0.176 \approx \lambda^* = 0.177.$$

The multiplier predicted the change almost exactly.

5. Second-Order Conditions

In previous classes, we classified critical points using the Hessian. For constrained problems, there is an analogous second-order condition involving the **bordered Hessian**. As is the “normal” Hessian, the bordered Hessian tells us whether the FOC solution is a maximum or a minimum of

the constrained problem.

However, you may have noticed that we have found the solution and did not check for the SOC's in last section's example. Why not? For utility maximization and cost minimization, the correct answer is almost always assured by the *shape* of the objective function in the first place, and we can confirm this without any matrix computation (yay!). The relevant concept is **quasi-concavity**.

Definition. A function $f(x, y)$ is **strictly quasi-concave** if, for any two distinct points A and B with $f(A) \geq f(B)$, every weighted average $\alpha A + (1-\alpha)B$ (with $0 < \alpha < 1$) satisfies $f(\alpha A + (1-\alpha)B) > f(B)$. In words: averages are strictly preferred to the worse of the two extremes.

Applied to a utility function, this means **strictly convex indifference curves** (bowed toward the origin). Applied to a production function, it means **strictly convex isoquants**.

If the objective function is strictly quasi-concave and the budget constraint is linear, any interior point satisfying the FOC is automatically a constrained maximum. The SOC is satisfied by the curvature of the objective function, not by any computation.

This definition is easiest to understand visually. The left panel of Figure 1 shows what this look like. In the left panel, the indifference curve is strictly quasi-concave: it bows smoothly toward the origin, with no bends the other way. Take any two bundles A and B (equally good, since they sit on the same indifference curve) and average them: the average bundle lands strictly above the curve, on a *better* indifference curve. Technically, it does not need to be the *exact* midpoint; any weighted average (in this context, called a **convex combination**) of A and B will do. The consumer strictly prefers the average bundle.

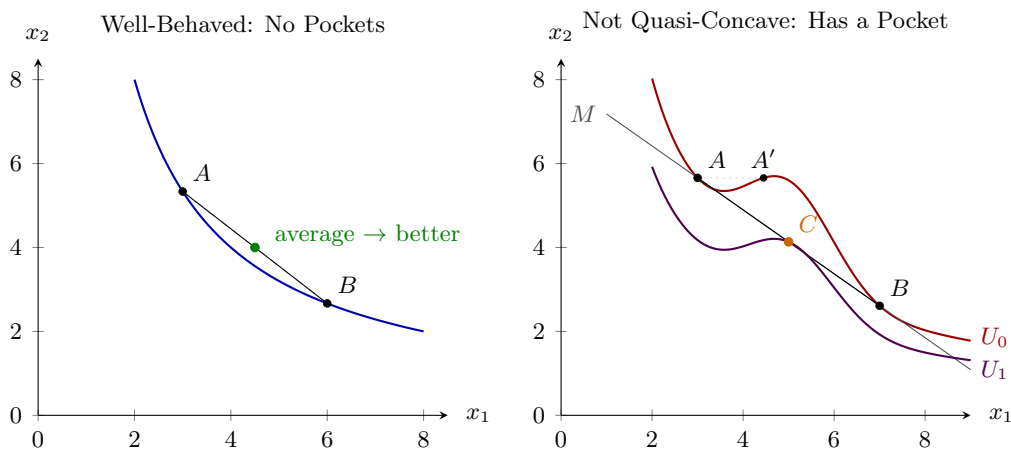


Figure 1: Strictly Quasi-Concave and Non-Quasi-Concave Indifference Curves.

Notes: Both panels plot an indifference curve U_0 through two equally-good bundles A and B . Left: the dashed chord marks the weighted averages of A and B as the weight α varies; the marked point (the 50-50 average) lies above the curve, on a strictly better indifference curve — quasi-concavity holds. Right: the curve dips inward then bulges outward, creating a pocket, so quasi-concavity fails. A' sits on U_0 at the same x_2 as A but with more x_1 , at no gain in utility. M is a budget line passing through A and B ; its midpoint C is affordable but lies on a strictly lower indifference curve, U_1 , not on U_0 — the FOC's tangency condition cannot distinguish a point like C from a genuine optimum.

It may not be immediately obvious why that is important, and why the tangency condition from the FOC is automatically the best point on the budget line, so it is important (and easier) to understand what can go wrong if the curve is *not* strictly quasi-concave. The right panel shows this scenario and there are two issues worth highlighting.

First, when the indifference curve bows inward for a while, creating a “pocket” or bulge, somewhere inside the pocket there is a third bundle, A' , with the *exact same* x_2 as A but strictly *more* x_1 ; and, since it sits on the same indifference curve, the consumer is no better off for having that extra x_1 . That alone should make you suspicious of a curve shaped this way.

Second, from the optimization perspective, a curve that is not strictly quasi-concave can lead to suboptimal solutions. Suppose the budget line, M , happens to pass through both A and B — both bundles are exactly affordable. Every bundle on the segment between them is affordable too, including the midpoint C . But C does not sit on the indifference curve through A and B at all; it sits on a strictly *lower* one, U_1 . A consumer who moves from A or B toward C along the budget line spends the same income and ends up with *less* utility. The ordinary tangency condition from the FOC cannot tell the difference between a point like C and a genuine optimum: it might just as easily land inside the pocket, at a point that is worse than nearby alternatives on either side.

Practical rule for this course

In every constrained optimization problem we will solve in this course, one of the two functions (the utility or production function) is strictly quasi-concave and the other (the budget or cost constraint) is linear. Therefore, if you find a solution that satisfies the FOC, you can be confident it is the correct constrained optimum.³ There are two exceptions to this, not because we will need to check the Hessian in these cases, but because the whole use of the Lagrangian method is not appropriate.

Perfect Complements. If the objective function is not differentiable, the Lagrangian method does not apply. This is the case of utility and production functions with kinks, such as perfect complements, also known as “fixed proportions” or Leontief functions.

Consider the following utility function:

$$U(x_1, x_2) = \min(\alpha x_1, \beta x_2), \quad \text{subject to } p_1 x_1 + p_2 x_2 = m.$$

Away from the kink, one of the two goods is not binding: if $\alpha x_1 > \beta x_2$, say, then $U = \beta x_2$ regardless of x_1 — the consumer could reduce x_1 without losing any utility at all. No rational consumer would spend money that way. So the optimal bundle must sit exactly on the kink, where the two terms are equal:

$$\alpha x_1 = \beta x_2 \implies x_2 = \frac{\alpha}{\beta} x_1$$

³This is by design, as we generally choose functional forms that are quasi-concave. In applied work, we sometimes need to check for SOC's when the functional form is unknown (we are estimating it using data), when there are multiple constraints or inequality constraints.

This is not a first-order condition, as the function has no derivative at the kink. It is simply the geometric fact that the best affordable bundle is always at a corner of the “L”, where the fixed proportions hold. Substitute directly into the budget constraint:

$$p_1 x_1 + p_2 \left(\frac{\alpha}{\beta} x_1 \right) = m \implies x_1^* = \frac{\beta m}{\beta p_1 + \alpha p_2}, \quad x_2^* = \frac{\alpha m}{\beta p_1 + \alpha p_2}$$

Thus, the Lagrangian and any derivatives are not required.

Perfect Substitutes. This is the case where the consumer is always willing to trade the two goods at the same rate, no matter how much of each she already has. The MRS never diminishes, and the total utility is simply the sum of the utilities from each good.⁴

Consider the following utility function:

$$U(x_1, x_2) = \alpha x_1 + \beta x_2, \quad \text{subject to } p_1 x_1 + p_2 x_2 = m.$$

Mathematically, the indifference curves are straight lines with constant slope $-\alpha/\beta$, so the FOC’s tangency condition is *never* satisfied. The optimum is always at a corner of the budget line (hence we call it a **corner solution**), where the consumer spends all her money on one good or the other.

In this case, the consumer compares the two goods and simply spends their entire budget on the one that delivers more utility *per dollar*.⁵ That is, the optimum is determined by comparing α/p_1 to β/p_2 , exactly as in Section 3:

$$\frac{\alpha}{p_1} > \frac{\beta}{p_2} \implies x_1^* = \frac{m}{p_1}, \quad x_2^* = 0 \qquad \frac{\alpha}{p_1} < \frac{\beta}{p_2} \implies x_1^* = 0, \quad x_2^* = \frac{m}{p_2}$$

Finally, the special case $\alpha = \beta$ is called **perfect substitutes**, where the two goods are completely interchangeable one-for-one. Think of two brands of bottled water with identical properties where you just buy the cheapest one.

⁴A sum between the two terms is not enough on its own for this to hold — the terms also need to be *linear*, so that each contributes a constant marginal utility rather than a diminishing one. For example, $U = \ln x_1 + \ln x_2$ is also a sum of two terms, but each term’s marginal utility diminishes with its own good ($d(\ln x_i)/dx_i = 1/x_i$), so the indifference curves are still curved and the ordinary FOC approach applies and thus it is not a case of perfect substitutes.

⁵If $\alpha/p_1 = \beta/p_2$ exactly, the indifference lines have the same slope as the budget line, and every bundle on the budget line is equally good, so there’s no unique solution.

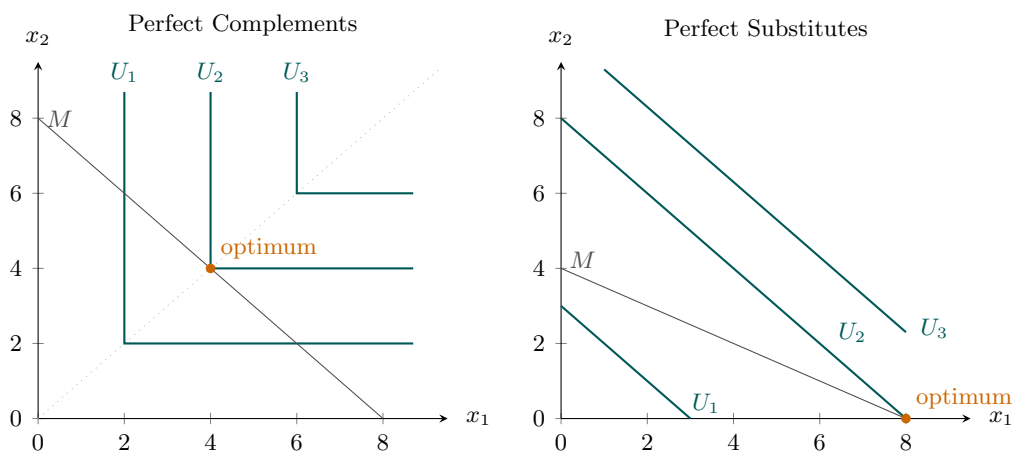


Figure 2: Perfect Complements and Perfect Substitutes.

Notes: M is the budget line in both panels. Left: kinked (L-shaped) indifference curves for $U = \min(\alpha x_1, \beta x_2)$. The dotted ray marks the kink locus $\alpha x_1 = \beta x_2$; the optimum is always on this ray, wherever it meets M (U_2 here) — U_1 is affordable but worse, U_3 is better but unaffordable. Right: straight-line indifference curves for $U = \alpha x_1 + \beta x_2$, with constant slope $-\alpha/\beta$. The highest affordable line (U_2) touches M only at a corner, where the consumer spends her entire income on one good; U_1 is affordable but worse, U_3 is better but unaffordable.

Practice Problems

1. A consumer has utility $U(x_1, x_2) = x_1^{0.25}x_2^{0.75}$, prices $p_1 = 4$, $p_2 = 1$, and income $m = 40$.
 - (a) Write the Lagrangian and the three FOC.
 - (b) Derive the MRS condition and use it to express x_1 in terms of x_2 .
 - (c) Solve for (x_1^*, x_2^*) and compute U^* .
 - (d) Find λ^* and interpret it as the marginal utility of income.
2. A consumer has utility $U(x_1, x_2) = \ln x_1 + 2 \ln x_2$, facing the budget constraint $p_1x_1 + p_2x_2 = m$. Although this is a *sum* of two terms, it is not the perfect-substitutes case from the previous section: each term is concave (a log), not linear, so the indifference curves are still smooth and bowed toward the origin, and the ordinary FOC approach applies.
 - (a) Write the Lagrangian and the three FOC. (Hint: recall that the derivative of $k \ln x$ is k/x .) Solve each FOC for λ and set the two expressions equal, then combine the resulting MRS condition with the budget constraint to solve for x_1^* and x_2^* in terms of p_1 , p_2 , and m .
 - (b) What fraction of income does the consumer spend on each good? How does this relate to the coefficients 1 and 2 in the utility function? (Hint: $\ln x_1 + 2 \ln x_2$ represents the same preferences as $x_1x_2^2$ — the coefficients play the same role as the exponents in a Cobb-Douglas utility function.)
 - (c) Find λ^* and interpret it as the marginal utility of income.
3. A firm must produce exactly $\bar{Q}$ units using labor L and capital K . Its production function is $Q = K^{0.5}L^{0.5}$, and input prices are w (wage) and r (rental rate).
 - (a) Set up the cost-minimization Lagrangian: minimize $C = wL + rK$ subject to $K^{0.5}L^{0.5} = \bar{Q}$, and write the three FOC.
 - (b) Solve each FOC for λ and set $MP_L/w = MP_K/r$ to eliminate λ .
 - (c) Use this condition to express K in terms of L .
 - (d) Substitute into the constraint to solve for L^* and K^* in terms of w , r , and $\bar{Q}$.
 - (e) Compute the minimum cost C^* in terms of w , r , and $\bar{Q}$.
 - (f) Interpret λ^* (the shadow price of the output constraint).
4. Explain in words and graphically, without any algebra, why the consumer's optimal bundle is at the point where the indifference curve is tangent to the budget line. Why can't the optimum be at an intersection (where the indifference curve crosses the budget line)? Sketch a budget line together with a few indifference curves to illustrate your answer.