Microeconomics

Chapter 4

Cost minimization

Fall 2024

Cost minimization

This chapter studies the behavior of a **cost minimizing** firm. This behavior is of interest for at least two reasons:

First, independent of how much output y a firm decides to produce, it should always aim to produce its output against minimum costs $\mathbf{wx}$. That is, a firm cannot maximize profits without minimizing costs.

Second, the analysis of cost minimization introduces us to the practice of **constrained optimization**.

Constrained optimization

Consider the problem of producing a certain level of output y against minimum costs:

$$\min_{\mathbf{x}} \mathbf{wx},$$
$$\text{such that } f(\mathbf{x}) = y.$$

In other words: the firm knows it wants to produce y, and then asks how it can produce that y against the lowest costs.

The method of Lagrange

First, write down the Lagrangian,

$$\mathcal{L} = \mathbf{wx} - \lambda(f(\mathbf{x}) - y).$$

Second, differentiate $\mathcal{L}$ wrt each endogenous variable: x_i for $i = 1, ..., n$ and λ. The FOCs for an interior solution $\mathbf{x}^*$ set these derivatives to zero,

$$\begin{aligned} w_1 - \lambda \frac{\partial f(\mathbf{x})}{\partial x_1} &= 0, \\ w_2 - \lambda \frac{\partial f(\mathbf{x})}{\partial x_2} &= 0, \\ &\vdots \\ w_n - \lambda \frac{\partial f(\mathbf{x})}{\partial x_n} &= 0, \\ f(\mathbf{x}) - y &= 0. \end{aligned}$$

Third, since we have $n + 1$ unknown endogenous variables (x_i for $i = 1, ..., n$ and λ) and $n + 1$ FOCs, we can solve for the endogenous variables in terms of the exogenous variables ($\mathbf{w}$ and y).

The method of Lagrange

Note that we can write the first n FOCs more efficiently:

$$\mathbf{w} = \lambda \mathbf{D} f(\mathbf{x}),$$

where $\mathbf{D}f(\mathbf{x})$ is the gradient of $f(\mathbf{x})$, and $\mathbf{w}$ is the gradient of $\mathbf{wx}$.

The method of Lagrange with two inputs

First, write down the Lagrangian,

$$\mathcal{L} = w_1 x_1 + w_2 x_2 - \lambda(f(\mathbf{x}) - y).$$

Second, differentiate $\mathcal{L}$ wrt each endogenous variable: x_1, x_2 and λ. The FOCs for an interior solution $\mathbf{x}^*$ set these derivatives to zero,

$$w_1 - \lambda \frac{\partial f(\mathbf{x})}{\partial x_1} = 0,$$

$$w_2 - \lambda \frac{\partial f(\mathbf{x})}{\partial x_2} = 0,$$

$$f(\mathbf{x}) - y = 0.$$

Third, since we have 3 unknown endogenous variables (x_1, x_2 and λ) and 3 FOCs, we can solve for the endogenous variables in terms of the exogenous variables (w_1, w_2 and y).

The method of Lagrange with two inputs

Dividing the first two FOCs by each other gives us the following optimality condition:

$$\frac{w_1}{w_2} = \frac{\partial f(\mathbf{x}) / \partial x_1}{\partial f(\mathbf{x}) / \partial x_2}.$$

We will combine this optimality condition with a graphical analysis of the Lagrange method to introduce the SOC and develop an economic intuition for the method more generally.

The method of Lagrange graphically

A firm's cost is equal to:

$$c = w_1 x_1 + w_2 x_2.$$

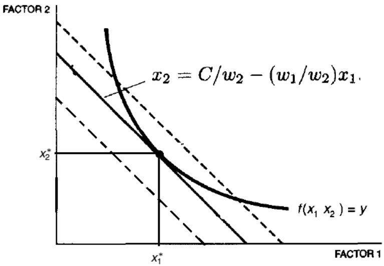
The **isocost** line is the level set for a single value of cost C:

$$L(C) = \{(x_1, x_2) : x_2 = \left(\frac{C}{w_2}\right) - \left(\frac{w_1}{w_2}\right) x_1\}.$$

Hence, the isocost line reflects all combinations of (x_1, x_2) that cost C.

The intercept of the isocost line $\left(\frac{C}{w_2}\right)$ gives the costs in terms of the input price of x_2. Note that a lower intercept implies lower costs (price w_2 is fixed).

The method of Lagrange graphically



The cost minimizing firm wants to find a **point on the isoquant with the minimal costs**: this is a point where the intercept of the isocost line $\left(\frac{C}{w_2}\right)$ is minimal. This point $\mathbf{x}^*$ is characterized by the slopes of the two lines being equal, which is the optimality condition we have seen before:

$$\frac{w_1}{w_2} = \frac{\partial f(\mathbf{x})/\,\partial x_1}{\partial f(\mathbf{x})/\,\partial x_2}.$$

Second-order condition

The SOC for cost-minimization is that:

$$V(y) \text{ is convex} \leftrightarrow f(\mathbf{x}) \text{ is concave} \leftrightarrow \mathbf{h}^T \mathbf{D}^2 f(\mathbf{x}) \mathbf{h} \leq 0.$$

With this condition we can be certain that the isoquant is always weakly above the isocost line. This is what we need for $\mathbf{x}^*$ to be cost-minimizing.

This guarantees that any change in inputs $\mathbf{x}$ that keeps output constant (that is, a change along the isoquant) must result in weakly higher costs (that is, an isocost line related to a weakly higher C).

The method of Lagrange: economic intuition

Recall that the optimality condition is:

$$\frac{w_1}{w_2} = \frac{\partial f(\mathbf{x})/\ \partial x_1}{\partial f(\mathbf{x})/\ \partial x_2}.$$

The RHS is the **technical rate of substitution** ($\frac{\partial x_2(x_1)}{\partial x_1} = -\frac{\partial f(\mathbf{x})/\ \partial x_1}{\partial f(\mathbf{x})/\ \partial x_2}$): when x_1 increases, how much does x_2 need to decrease as to keep output constant.

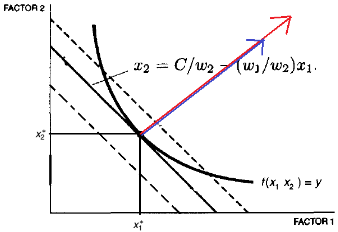
The LHS is the **economic rate of substitution** ($\frac{\partial x_2(x_1)}{\partial x_1} = -\frac{w_1}{w_2}$): when x_1 increases, how much does x_2 need to decrease as to keep costs constant.

Hence, the optimality condition tells us that at $\mathbf{x}^*$ the economic and technical rate of substitution need to be equal. Imagine that they are not:

$$\frac{w_1}{w_2} = \frac{2}{1} \neq \frac{1}{1} = \frac{\partial f(\mathbf{x})/\ \partial x_1}{\partial f(\mathbf{x})/\ \partial x_2},$$

then we can use one unit less of x_1 and one unit more of x_2, so that output remains unchanged but costs have gone down. This cannot be optimal.

The method of Lagrange: what about λ?



Then why does the method require λ? **The gradient of a function is perpendicular to the level set of that function**. Hence, $\mathbf{D}f(\mathbf{x}^*)$ is perpendicular to the isoquant and $\mathbf{w}$ is perpendicular to the isocost line. Since the slopes of the isoquant and the isocost line must be equal at $\mathbf{x}^*$, both gradients must "point in the same direction" at $\mathbf{x}^*$. However, they are not necessarily of "equal length" at $\mathbf{x}^*$. Hence, we multiply $\mathbf{D}f(\mathbf{x})$ by a scalar λ:

$$\mathbf{w} = \lambda \mathbf{D}f(\mathbf{x}).$$

These are simply the FOCs of the Lagrangian. It turns out that λ also has a useful economic interpretation. More on this later.

The conditional factor demand function and cost function

Conditional factor demand function $\mathbf{x}(\mathbf{w}, y)$: a function that gives us the optimal choice of inputs as a function of the input prices and desired output level.

How to get this function? From the FOCs of the Langrangian we can write $\mathbf{x}$ in terms of $(\mathbf{w}, y)$.

Cost function $c(\mathbf{w}, y)$: a function that gives us the minimal costs for producing y units of output against input prices $\mathbf{w}$.

How to get this function? Substitute $\mathbf{x}(\mathbf{w}, y)$ into $\mathbf{wx} = \mathbf{wx}(\mathbf{w}, y) = c(\mathbf{w}, y)$.

Exercise

Derive the conditional factor demand functions for the following cost minimization problem:

$$\min_{x_1, x_2} w_1 x_1 + w_2 x_2 \quad \text{subject to } x_1^{\alpha} x_2^{1-\alpha} = y.$$

Exercise

Derive the conditional factor demand functions for the following cost minimization problem:

$$\min_{x_1, x_2} w_1 x_1 + w_2 x_2 \quad \text{subject to } (x_1^{\rho} + x_2^{\rho})^{\frac{1}{\rho}} = y.$$

Exercise

Derive the conditional factor demand functions for the following cost minimization problem:

$$\min_{x_1, x_2} w_1 x_1 + w_2 x_2 \quad \text{subject to } x_1 + x_2 = y.$$

Hint: draw the isoquant and isocost line.

The Lagrange multiplier

How to interpret the Lagrange multiplier λ? It turns out that this endogenous variable has a useful economic interpretation.

The Lagrange multiplier λ measures how the optimal solution to the constrained optimization problem changes when the constraint is relaxed.

When we apply this interpretation of λ to the cost minimization problem, the optimal solution is the cost function $c(\mathbf{w}, y)$ and the constraint is relaxed if we increase the amount of production y.

Hence, in this setting the Lagrange multiplier measures how the costs change, $\Delta c(\mathbf{w}, y)$, when we increase the production, Δy, so $\lambda = \frac{\Delta c(\mathbf{w},y)}{\Delta y} = \frac{\partial c(\mathbf{w},y)}{\partial y}$, which are the **marginal costs**.

The Lagrange multiplier

The proof for this interpretation of λ follows from the envelope theorem.

Consider the Lagrangian with two inputs,

$$\mathcal{L}(\mathbf{w}, y, \mathbf{x}, \lambda) = w_1 x_1 + w_2 x_2 - \lambda(f(\mathbf{x}) - y).$$

First, note that:

$$\frac{\partial \mathcal{L}(\mathbf{w}, y, \mathbf{x}, \lambda)}{\partial y} = \lambda.$$

Second, substitute the conditional factor demand functions $\mathbf{x}(\mathbf{w}, y)$ and the Lagrange multiplier $\lambda(\mathbf{w}, y)$ into the Lagrangian to obtain the Lagrangian evaluated at the optimal point: $\mathcal{L}(\mathbf{w}, y, \mathbf{x}(\mathbf{w}, y), \lambda(\mathbf{w}, y)) = \mathcal{L}(\mathbf{w}, y)$. It turns out, this is equal to:

$$\begin{aligned}
\mathcal{L}(\mathbf{w}, y) =& w_1 x_1(\mathbf{w}, y) + w_2 x_2(\mathbf{w}, y) - \lambda(\mathbf{w}, y)\Big(f(\mathbf{x}(\mathbf{w}, y)) - y\Big), \\
=& w_1 x_1(\mathbf{w}, y) + w_2 x_2(\mathbf{w}, y), \\
=& \mathbf{w}\mathbf{x}(\mathbf{w}, y), \\
=& c(\mathbf{w}, y).
\end{aligned}$$

The Lagrange multiplier

Third, use the logic of the envelope theorem to show that at the optimal point:

$$\begin{aligned}\frac{\partial \mathcal{L}(\mathbf{w},y)}{\partial y} =& \underbrace{\frac{\partial \mathcal{L}(\cdot)}{\partial y}}_{\text{direct effect}} + \underbrace{\frac{\partial \mathcal{L}(\cdot)}{\partial x_1}\frac{\partial x_1(\cdot)}{\partial y} + \frac{\partial \mathcal{L}(\cdot)}{\partial x_2}\frac{\partial x_2(\cdot)}{\partial y} + \frac{\partial \mathcal{L}(\cdot)}{\partial \lambda}\frac{\partial \lambda(\cdot)}{\partial y}}_{\text{indirect effect}}, \\ =& \frac{\partial \mathcal{L}(\cdot)}{\partial y}, \\ =& \lambda(\mathbf{w},y),\end{aligned}$$

as the indirect effects are zero because of the FOCs of the Lagrangian.

Since $\mathcal{L}(\mathbf{w},y) = c(\mathbf{w},y)$, we conclude that:

$$\frac{\partial \mathcal{L}(\mathbf{w},y)}{\partial y} = \frac{\partial c(\mathbf{w},y)}{\partial y} = \lambda(\mathbf{w},y).$$

Exercise

Consider again the following cost minimization problem,

$$\min_{x_1, x_2} w_1 x_1 + w_2 x_2 \quad \text{subject to } x_1^{\alpha} x_2^{1-\alpha} = y.$$

1. On top of the conditional factor demand functions derived in the exercise above, also derive the Lagrange multiplier $\lambda(\mathbf{w}, y)$.

2. Now use the conditional factor demand functions to derive the cost function $c(\mathbf{w}, y)$. Then show that:

$$\frac{\partial c(\mathbf{w}, y)}{\partial y} = \lambda(\mathbf{w}, y).$$

3. Provide an economic interpretation to $\lambda(\mathbf{w}, y)$.

Weak Axiom of Cost Minimization

Similar to the WAPM, consider you observe the following data for a firm: a row vector of input prices $\mathbf{w}^t$ and a column vector of input and output levels $\mathbf{x}^t$ and y^t across time $t = 1, ..., T$. Hence, we have data for a firm over time t.

WACM: a necessary condition for cost minimization is that,

$$\underbrace{\mathbf{w}^t\mathbf{x}^t}_{\text{actual costs}} \leq \underbrace{\mathbf{w}^t\mathbf{x}^s}_{\text{potential costs}}, \quad \forall\, s, t \text{ with } y^s \geq y^t.$$

If the firm minimizes costs, then the actual costs of the observed choice of inputs $\mathbf{x}^t$ should be no greater than the potential costs at any other level of inputs that the firm could have chosen and would have produced at least as much output. We do not know all such inputs $\mathbf{x}$, but we do know some, namely $\mathbf{x}^s$ with $y^s \geq y^t$.

Hence, only with data on a firm's input prices $\mathbf{w}^t$ and input and output levels $\mathbf{x}^t$ and y^t across time t you may conclude that the firm makes choices that do not minimize costs.

Homework exercises

Exercises: 4.1, 4.7, and exercises on the slides