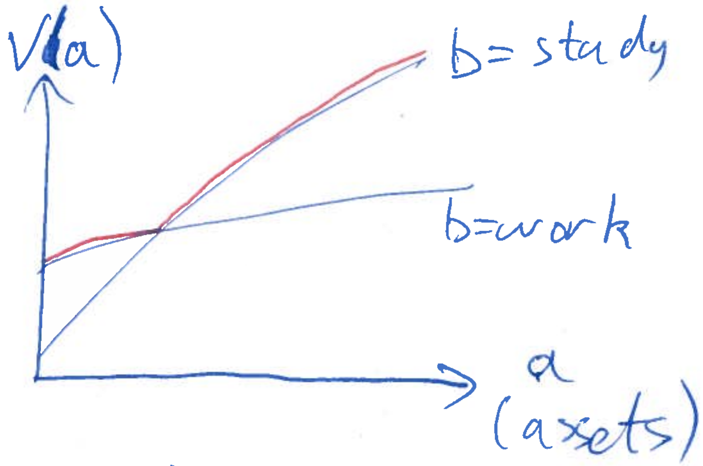
2.3 Envelope theorem etc.

Economically speaking:

$$\underbrace{\underbrace{V(a)}_{\text{state variable}}}_{\text{value function}} = \max_{\underbrace{b}_{\text{choice variable}}} \underbrace{v(a, b)}_{\text{objective function}} = v(a, \underbrace{b(a)}_{\text{policy function}})$$

e.g.



Goal: study relationship between marginal values $V'(a)$ and optimal choices $b(a)$.

The red curve envelopes ("surrounds") the blue lines.
The red curve is called the "upper envelope"

High school geometry: if two curves touch without crossing, then they are tangent (i.e. have the same scope, and are parralel).

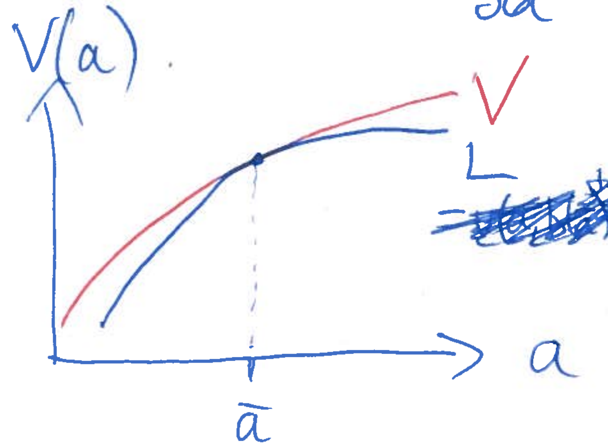
Theorem 2.1 (Envelope theorem)

Let $v: \underset{\substack{a \\ \text{(states)}}}{\mathbb{R}^n} \times \underset{\substack{b \\ \text{(choices)}}}{\mathbb{R}^m} \to \mathbb{R}$ be a differentiable,

$V(a) = \max_{b \in \mathbb{R}^m} v(a, b)$ be its value function and $b(a)$ be its policy function.

If V is a differentiable function, then
$$V'(a) = \left.\frac{\partial v(a, b)}{\partial a}\right|_{b = b(a)}.$$

Proof:



(Rupert Murdoch)

(James Murdoch) "lazy"

$b = \#$ lies

$a =$ power balance

Let $L(a) = v(a, \underbrace{b(\bar{a})}_{\text{"final words"}})$.

Since $b(\bar{a})$ is inferior to $b(a)$ at state a, (i.e. $v(a, b(\bar{a})) \le v(a, b(a))$) we know $L(a) \le V(a)$, for all a.

We also know $L(\bar{a}) = V(\bar{a})$. Therefore $V(a) - L(a)$ is minimised at $\bar{a}$. At the minimum, the first-order condition $V'(\bar{a}) - L'(\bar{a}) = 0$ holds, $\Longleftrightarrow$ $V'(\bar{a}) = L'(\bar{a})$.

where $L'(\bar{a}) = \left.\frac{\partial v(a, b)}{\partial a}\right|_{b = b(\bar{a})}$. $\square$

Proof (chain rule):

$V(a) = v(a, b(a))$.

Chain rule says:

$$V'(a) = \underbrace{\left.\frac{\partial v(a,b)}{\partial a}\right|_{b=b(a)}}_{\text{direct effect}} + \underbrace{\overbrace{\left.\frac{\partial v(a,b)}{\partial b}\right|_{b=b(a)}}^{\text{FOC} \Rightarrow 0} b'(a)}_{\text{indirect effect}}$$

Detour: vector calculus

Consider $h(x) = f(g(x))$,

where $g: \mathbb{R} \to \mathbb{R}^2$, $f: \mathbb{R}^2 \to \mathbb{R}$

Chain rule: $h'(x) = f'(g(x))\, g'(x)$.

$$= \begin{bmatrix} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial x} \\ \frac{\partial g_2}{\partial x} \end{bmatrix}$$

$$= \frac{\partial f}{\partial y_1}\frac{\partial g_1}{\partial x} + \frac{\partial f}{\partial y_2}\frac{\partial g_2}{\partial x}.$$

Link to envelope theorem:

$x = a$, $h(x) = V(a)$, $g(x) = \begin{bmatrix} a \\ b(a) \end{bmatrix}$,

$y = \begin{bmatrix} a \\ b \end{bmatrix}$, $f(y) = v(a, b)$, $f(g(x)) = v(a, b(a))$

Since $b(a)$ is an optimal choice at a, we know the first-order condition

$$\left.\frac{\partial v(a,b)}{\partial b}\right|_{b=b(a)} = 0.$$

Therefore the second term vanishes and the formula we wanted is left over. $\square$

Example: ℓ workers
w wage

Profits:

$$\pi(w) = \max_{\ell} 10\sqrt{\ell} - w\ell.$$

Want to calculate $\pi'(w)$.

With envelope theorem:

$$\pi'(w) = \left[\frac{\partial}{\partial w}[10\sqrt{\ell} - w\ell]\right]_{\ell=\ell(w)}$$

$$= [-\ell]_{\ell=\ell(w)}$$

$$= -\ell(w).$$

Without the envelope theorem:

~~Refo~~ Reformulate $\pi(w) = 10\sqrt{\ell(w)} - w\,\ell(w)$ and differentiate.

$$\pi'(w) = 10 \cdot \frac{1}{2}\frac{1}{\sqrt{\ell(w)}} \cdot \ell'(w) - \ell(w) - w\ell'(w)$$

$$= 5\frac{\ell'(w)}{\sqrt{\ell(w)}} - \ell(w) - w\ell'(w).$$

What a mess!

To calculate $\ell(w)$, we take the FOC

$$\frac{\partial}{\partial \ell}[10\sqrt{\ell} - w\ell] = 0$$

$$\Leftrightarrow 10 \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{\ell}} - w = 0$$

$$\Leftrightarrow \frac{5}{\sqrt{\ell}} = w \quad \Leftrightarrow \sqrt{\ell} = \frac{5}{w}$$

$$\Leftrightarrow \ell(w) = \frac{25}{w^2}.$$

$$\ell'(w) = -\frac{50}{w^3}.$$

Substituting into

$$\pi'(w) = 5\frac{\ell'(w)}{\sqrt{\ell(w)}} - \ell(w) - w\,\ell'(w)$$

$$= 5\frac{\ell'(w)}{\left(\frac{5}{w}\right)} - \frac{25}{w^2} - w\,\ell'(w)$$

$$= \cancel{w\,\ell'(w)} - \frac{25}{w^2} - \cancel{w\,\ell'(w)}$$

$$= -\frac{25}{w^2} = -\ell(w). \qquad \square$$

Recall the firm's problem

$$\pi(p;w) = \max_{x \in \mathbb{R}^{N-1}_+} pf(x) - w \cdot x$$

Applying the envelope theorem gives:

$$\frac{\partial \pi(p;w)}{\partial p} = \left[\frac{\partial}{\partial p}\{pf(x) - w \cdot x\}\right]_{x=x(p;w)}$$

$$= [f(x)]_{x=x(p;w)}$$

$$= f(x(p;w)) = \underbrace{y(p;w)}_{\text{output quantity}}.$$

$$\frac{\partial \pi(p;w)}{\partial w_i} = \left[\frac{\partial}{\partial w_i}\{pf(x) - w \cdot x\}\right]_{x=x(p;w)}$$

$$= [-x_i]_{x=x(p;w)}$$

$$= -x_i(p;w).$$

Differentiating again gives

$$\frac{\partial^2 \pi(p;w)}{\partial p^2} = \frac{\partial y(p;w)}{\partial p}$$

eg: i = doctors, j = nurses

$$\underbrace{\frac{\partial^2 \pi(p;w)}{\partial w_i \partial w_j}}_{\text{Assuming } \pi \text{ is twice differentiable.}} = -\frac{\partial x_i(p;w)}{\partial w_j} = -\frac{\partial x_j(p;w)}{\partial w_i}$$

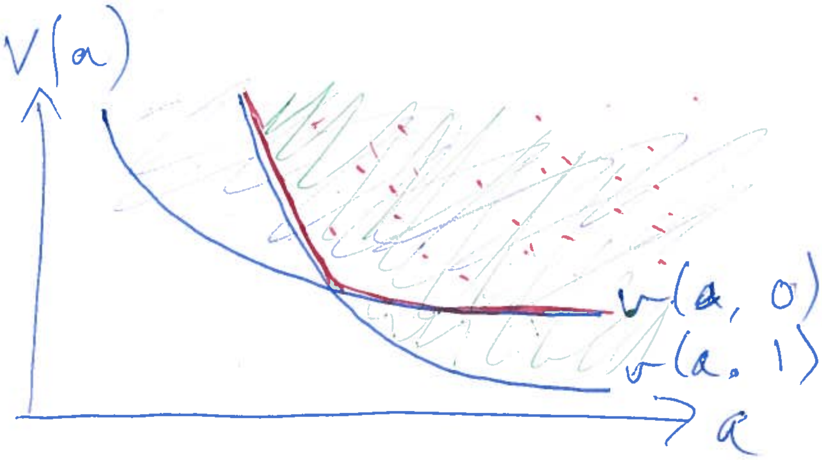
Assuming π is twice differentiable.

Theorem 2.2

Suppose $V(a) = \max_b v(a,b)$ where each $v(\cdot, b)$ is a convex function. Then V is a convex function.

Proof

Geometric idea:



Assume convex hyper($v(\cdot,0)$)

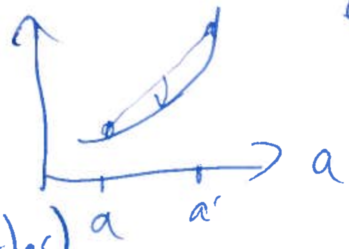
convex hyper($v(\cdot,1)$)

Want to prove convex hyper(V)

Since $\cap$ $=$, and the intersection of convex sets is convex, we conclude (hyper(V)) is a convex set. Therefore V is a convex function. $\square$

Proof (algebraic):

We would like to show



$$\underbrace{tV(a) + (1-t)V(a')}_{\text{line}} \geq \underbrace{V(ta + (1-t)a')}_{\text{curve}},$$

$t V(a) + (1-t)\, V(a')$

$= t\, v(a, b(a)) + (1-t)\, v(a', b(a'))$

$\geq t\, v(a, \underbrace{b(ta + (1-t)a')}_{\text{right choice for wrong state}}) + (1-t)\, v(a', b(a'))$

$\geq t\, v(a, b(ta + (1-t)a')) + (1-t)\, v(a', \underbrace{b(ta + (1-t)a')}_{\text{right choice for the same wrong state}})$

$\geq v(ta + (1-t)a', b(ta + (1-t)a'))$

since both points are on the same curve, and we assumed each curve (one for each choice) is a ~~function~~ convex function

$= V(ta + (1-t)a')$. $\square$

Theorem 2.3 For every production f, the firm's profit function π is a convex function. Hence if π is smooth, then

$$\frac{\partial y(p;w)}{\partial p} \geq 0 \quad \text{and} \quad \frac{\partial x_i(p;w)}{\partial w_i} \leq 0$$

Proof Step 1: π is convex:

For each input choice $x \in \mathbb{R}^{N-1}_+$, the ~~fn~~ objective function

$$(p;w) \mapsto p f(x) - w \cdot x$$

is a linear (and hence convex) function.

Therefore, π is the upper envelope of a set of convex functions, so Theorem 2.2 establishes that π is a convex function

<u>Step 2</u>: By the envelope theorem, we previously established that

$$\frac{\partial \pi(p;w)}{\partial p} = y(p;w),$$

and

$$\frac{\partial^2 \pi(p;w)}{\partial p^2} = \frac{\partial y(p;w)}{\partial p}.$$

Since π is a convex function, LHS ≥ 0. So RHS ≥ 0, i.e.

$$\frac{\partial y(p;w)}{\partial p} \geq 0. \qquad \square$$

Note: only works if we diff'ate w.r.t. the same variable twice.

Lessons:

* envelope theorem – how to diff'ate value functions

* env. theorem told us: marginal values are related to optimal choices

* even without convexity assumptions on the production function, we established that the profit function is convex.

* We established the input choices are decreasing in their corresponding prices, and output is increasing in the output price.

2.4 Cost functions and Dynamic Programming

Recall

$$\pi(p; w) = \max_{x \in \mathbb{R}_+^{N-1}} pf(x) - w \cdot x.$$

$$\pi(p; w) = \max_{y \in \mathbb{R}_+} py - c(y; w)$$

Bellman equation

$$\text{where } \underbrace{c(y; w)}_{\text{cost function}} = \min_{x \in \mathbb{R}_+^{N-1}} w \cdot x \quad \text{s.t. } f(x) \geq y.$$

meet prod. target

Lemma 2.1 (Principle of Optimality)

$$\pi(p;w) = \max_{x\in\mathbb{R}_+^{N-1}} pf(x) - w\cdot x$$

$$= \max_{y\in\mathbb{R}_+,\, x\in\mathbb{R}_+^{N-1}} pf(x) - w\cdot x \quad \text{s.t. } f(x) = y$$

$$= \max_{y\in\mathbb{R}_+} \left(\max_{x\in\mathbb{R}_+^{N-1}} pf(x) - w\cdot x \quad \text{s.t. } f(x) = y \right)$$

special case of $\max_{a,b} f(a,b) = \max_a \max_b f(a,b)$

$$= \max_a g(a)$$

where $g(a) = \max_b f(a,b)$

$$= \max_{y\in\mathbb{R}_+} \left(\max_{x\in\mathbb{R}_+^{N-1}} py - w\cdot x \quad \text{s.t. } f(x) = y \right)$$

$$= \max_{y\in\mathbb{R}_+} py + \left(\max_{x\in\mathbb{R}_+^{N-1}} \cancel{y} - w\cdot x \quad \text{s.t. } f(x) = y \right)$$

$$= \max_{y\in\mathbb{R}_+} py - \left(\min_{x\in\mathbb{R}_+^{N-1}} w\cdot x \quad \text{s.t. } f(x) = y \right)$$

$$= \max_{y\in\mathbb{R}_+} py - c(y;w) \qquad \square$$

Theorem 2.4

$$p = \underbrace{\frac{\partial c(y;w)}{\partial y}}_{\text{marginal cost}} \Bigg|_{y = y(p;w)}$$

__Proof__ Consider the Bellman equation

$$\pi(p;w) = \max_{y \in \mathbb{R}_+} py - c(y;w).$$

The FOC w.r.t. y is

$$p = \left.\frac{\partial c(y;w)}{\partial y}\right|_{y=y(p;w)}$$

as required. $\square$

We can also apply the envelope theorem to the Bellman equation:

$$\frac{\partial \pi(p;w)}{\partial p} = \left[\frac{\partial}{\partial p}\{py - c(y;w)\}\right]_{y=y(p;w)}$$

$$= [y]_{y=y(p;w)}$$

$$= y(p;w)$$

and

$$\frac{\partial \pi(p;w)}{\partial w_i} = \left[\frac{\partial}{\partial w_i}\{py - c(y;w)\}\right]_{y=y(p;w)}$$

$$= \left[-\frac{\partial c(y;w)}{\partial w_i}\right]_{y=y(p;w)}.$$

2.5 Upper Envelopes with Constraints

$$V(a) = \max_b \; v(a,b)$$
$$\text{s.t. } w(a,b) \geq 0$$

add a constraint

Lagrangian: $\mathcal{L}(a,b,\lambda) = v(a,b) + \lambda w(a,b)$.

~~Theorem~~ It's often true that

$$V(a) = \min_\lambda \max_b \mathcal{L}(a,b,\lambda).$$

FOC w.r.t. b:

$$\left.\frac{\partial \mathcal{L}(a,b,\lambda)}{\partial b}\right|_{b=b(a),\, \lambda=\lambda(a)} = 0$$

$$\Leftrightarrow \left[\frac{\partial v(a,b)}{\partial b} + \lambda \frac{\partial w(a,b)}{\partial b}\right]_{b=b(a),\, \lambda=\lambda(a)}$$

Theorem 2.5 (Constrained Envelope Theorem)

If V, v, w, b, λ are all differentiable functions, and the constraint binds at a, then

$$V'(a) = \left[\frac{\partial v(a,b)}{\partial a} + \lambda \frac{\partial w(a,b)}{\partial a}\right]_{\substack{b=b(a),\\ \lambda=\lambda(a)}}$$

Proof (Chain rule proof).

$$V(a) = v(a, b(a)).$$

Lagrangian trick:

$$V(a) = v(a, b(a)) + \lambda(a) w(a, b(a))$$
$$= L(a, b(a), \lambda(a)).$$

Differentiating gives:

$$V'(a) = \left[\frac{\partial L(a,b,\lambda)}{\partial a} + \underbrace{\frac{\partial L(a,b,\lambda)}{\partial b}}_{=0} b'(a) + \frac{\partial L(a,b,\lambda)}{\partial \lambda} \cdot \lambda'(a) \right]_{b=b(a),\, \lambda=\lambda(a)}$$

By the FOC, the circled bit is 0. □