1 Solution: produce or learn?

The **optimal strategy** has a very simple form: there is a time $t^* \in \{1, \ldots, T\}$ such that (* denotes optimal quantities):

$$\begin{aligned} \text{before time } t^* &: \text{ only learn, i.e., } l^*(t) = 1 \text{ for } t < t^* \\ \text{from time } t^* &: \text{ only produce, i.e., } p^*(t) = 1 \text{ for } t \geq t^* \end{aligned}$$

This start-production time t^* can be explicitly determined (see proof below). Except possibly at time t^*, it is never optimal to split time between learning and producing ($l^*(t), p^*(t) \in (0,1)$).

The optimal strategy has the following interpretation:

$$\begin{aligned} \text{learn at full speed } (l^*(t) = 1) \quad &\text{if} \quad \text{present quality is less than future capacity} \\ \text{produce at full speed } (p^*(t) = 1) \quad &\text{if} \quad \text{present quality exceeds future capacity} \end{aligned}$$

Let's first explain this interpretation more precisely and then prove that this strategy is optimal.

1.1 Interpretation

Let's call $Q(t) := P(t) + mL(t)$ the *quality at time t*. Since the objective function is strictly increasing in $p(t)$ and $l(t)$, at optimality, $p^*(t) + l^*(t) = 1$. We can hence eliminate $l(t)$ by substituting $l(t) = 1 - p(t)$ into $Q(t)$ to get:

$$Q(t) \;=\; \sum_{s=1}^{t-1} p(s) + m \sum_{s=1}^{t-1} (1 - p(s)) \;=\; m(t-1) - (m-1) \sum_{s \leq t-1} p(s)$$

Note that $Q(1) = P(1) + mL(1) = 0$ since $p(0) := 0 =: l(0)$. Our problem is then reduced to:

$$\begin{aligned} \max_{p(t), t=1,\ldots,T} \quad & f(p) \;:=\; \sum_{t=1}^{T} p(t)\, Q(t) \\ \text{subject to} \quad & 0 \leq p(t) \leq 1, \quad t = 1, \ldots, T \end{aligned}$$

Taking partial derivatives of the objective function:

$$\frac{\partial f}{\partial p(t)} \;=\; Q(t) \;-\; (m-1) \sum_{s \geq t+1} p(s) \;=:\; Q(t) \;-\; \overline{Q}(t) \tag{1}$$

where the *marginal capacity-to-go* (from time $t+1$ to T) is defined as

$$\overline{Q}(t) \;:=\; (m-1) \sum_{s \geq t+1} p(s)$$

If the quality $Q(t)$ at the present time t is less than the marginal capacity-to-go $\overline{Q}(t)$, i.e., if $Q(t) < \overline{Q}(t)$, then

$$\frac{\partial f}{\partial p(t)} \quad < \quad 0$$

and hence a smaller $p(t)$ will increase the objective value f. This means the optimal production effort $p^*(t) = 0$, i.e., learn fast and build up quality before producing. Conversely if $Q(t) > \overline{Q}(t)$, then $\partial f \partial p(t) > 0$ and $p^*(t) = 1$, i.e., produce at full speed if present quality exceeds the marginal capacity-to-go.

1.2 Proof

To solve for the optimal $p^*(t)$ more explicitly, we note from (1) that

$$\frac{\partial f}{\partial p(t)} \quad = \quad m(t-1) \;-\; (m-1)\,(P - p(t)) \tag{2}$$

where $P := \sum_{t=1}^{T} p(t)$ is the *total production capability.* Our solution follows from the following properties.

It is optimal to initially learn and don't produce. In that case, the marginal production value increases rapidly at rate m, until eventually produce at full speed. More precisely, we have

Proposition 1 *Suppose $T \geq 2$. Then, at optimality,*

1. $p^*(1) = 0$.
2. *If $p^*(t) = 0$ then*

$$\frac{\partial f}{\partial p(t+1)} \;-\; \frac{\partial f}{\partial p(t)} \quad \geq \quad m$$

3. *There exists some time $t \leq T$ at which $p^*(t) = 1$.*

Proof. We have from (2), at optimality,

$$\frac{\partial f}{\partial p(1)} \quad = \quad -(m-1)\,(P^* - p^*(1)) \quad = \quad -(m-1)\sum_{s=2}^{T} p^*(s) \quad < \quad 0$$

where the last inequality holds as long as $T \geq 2$. Hence $p^*(1) = 0$.

The second claim follows directly from (2). To prove the third claim, suppose $p^*(s) = 0$ for $s = 1, 2, \ldots, t-1$. Then $P^* - p^*(t) = \sum_{s=t+1}^{T} p^*(s)$. We have from (2)

$$\frac{\partial f}{\partial p(t)} \;=\; m(t-1) - (m-1)\sum_{s=t+1}^{T} p^*(s)$$

Hence if $p^*(t) = 0$ for all $t = 1, \ldots, T$, then $\frac{\partial f}{\partial p(t)}(T) = m(T-1) > 0$. This means $p^*(T) = 1$, a contradiction. Hence there exists an $t \leq T$ such that $p^*(t) = 1$. ■

Proposition 1 says that there is a time t^* at which the optimal strategy $p^*(t)$ switches from 0 to 1. What are the optimal values of $p^*(t)$ after time t^*? We claim that $p^*(t) = 1$ for all $t \geq t^*$.

Proposition 2 *Suppose $T \geq 2$. If $p^*(t) = 0$ for $t = 1, \ldots, \tau - 1$ and $p^*(\tau) = 1$ for some τ, then $p^*(t) = 1$ for $t = \tau, \ldots T$.*

Proof. Since $p^*(t) = 0$ for $t = 1, \ldots, \tau - 1$ and $p^*(\tau) = 1$, we have from (2), for $t = \tau, \ldots, T$,

$$\frac{\partial f}{\partial p(t)} \;=\; m(t-1) - (m-1)\left(\sum_{s=\tau}^{T} p(s) \;-\; p^*(t)\right)$$

Compare this quantity at time τ and at any time $t > \tau$:

$$\frac{\partial f}{\partial p(t)} - \frac{\partial f}{\partial p(\tau)} \;=\; m(t-\tau) - (m-1)\left(p^*(\tau) - p^*(t)\right) \;>\; 0$$

where the last inequality follows because $t - \tau \geq 1$ but $p^*(\tau) - p^*(t) \leq 1$. Hence (noting that $p^*(\tau) = 1$)

$$\frac{\partial f}{\partial p(t)} \;>\; \frac{\partial f}{\partial p(\tau)} \;\geq\; 0$$

This implies $p^*(t) = 1$ for $t = \tau, \ldots T$. ■

The only remaining task is to determine the time t^* at which $p^*(t)$ switches from 0 to 1. Even though we don't know the value of $p^*(t^*)$, Propositions 1 and 2 imply that the values of $p^*(t)$ at other times are either 0 or 1:

$$p^*(t) \;=\; \begin{cases} 0 & \text{for} \quad t = 1, \ldots, t^* - 1 \\ 1 & \text{for} \quad t = t^* + 1, \ldots, T \end{cases}$$

Hence $P^* - p^*(t^*) := \sum_{s=1}^{T} p^*(s) - p^*(t^*) = T - t^*$. Hence, from (2), we have

$$\frac{\partial f}{\partial p(t^*)} \;=\; m(t^*-1) \;-\; (m-1)(T-t^*) \tag{3}$$

The time t^* is when $\frac{\partial f}{\partial p(t)}$ switches from negative to positive. Since t^* must be an integer, (3) may not be 0. Let s^* be the real number that satisfies $m(s^* - 1) - (m - 1)(T - s^*) = 0$, i.e.,

$$s^* \quad := \quad \frac{(T+1)m - T}{2m - 1} \tag{4}$$

Then $t^* = \lceil s^* \rceil$.

This completes the derivation of the optimal strategy.

Remark. How does the quality parameter m impact the start time t^*? From (4), we see that, as m changes from 1 to ∞, t^* changes from 1 to $(T+1)/2$. Hence, as quality becomes more important (larger m), one starts to produce later in order to build up quality (up to roughly half the horizon T).