

Universidad Nacional de Ingeniería  
Facultad de Ciencias  
Av. Túpac Amaru N° 210 Rimac  
Lima, Perú

Presentación teórica del tema

## La función de producción Cobb-Douglas

La función de producción: Cobb-Douglas

K. Fernández Huidobro*	2018228G
A. Berrospi C.†	20182720C
B. Torres Ayalas‡	201721339G
C. Aznarán Laos§	20162720C

21 de octubre del 2019

Supervisado por M.Sc. Clifford Torres Ponce¶

---

\*Email: [kfernandezh@uni.pe](mailto:kfernandezh@uni.pe)

†Email: [aaberrospic@uni.pe](mailto:aaberrospic@uni.pe)

‡Email: [btorresa@uni.pe](mailto:btorresa@uni.pe)

§Email: [caznaranl@uni.pe](mailto:caznaranl@uni.pe)

¶[ctorresp@uni.edu.pe](mailto:ctorresp@uni.edu.pe)

## Resumen

La función de producción Cobb-Douglas es un enfoque neoclásico para estimar la función de producción de un país y proyectar de esta manera su crecimiento económico esperado. Para representar las relaciones entre la producción obtenida se utiliza las variaciones de los insumos como el capital ( $K$ ) y el trabajo ( $L$ ), a los que más tarde se añadió la tecnología, llamada también productividad total de los factores ( $PTF$ ). Es una función de producción frecuentemente utilizada en Economía.

El origen de la función Cobb-Douglas se encuentra en la observación empírica de la distribución de la renta nacional total de Estados Unidos entre el capital y el trabajo. De acuerdo a lo que mostraban los datos, la distribución se mantenía relativamente constante a lo largo del tiempo. Concretamente el trabajo se llevaba un 70% y el capital un 30%. De esta forma, la función Cobb-Douglas representa una relación en donde las proporciones de trabajo y capital con respecto al producto total son constantes.

# Contenido

<b>1. Introducción</b>	<b>4</b>
1.1. Función de producción de Cobb-Douglas . . . . .	4
1.2. Función de producción . . . . .	4
<b>2. El proyecto</b>	<b>4</b>
2.1. Cobb-Douglas y la función de producción ACMS . . . . .	4
2.2. Breve descripción . . . . .	5
<b>3. Introducción</b>	<b>11</b>
<b>4. Un modelo de crecimiento a largo plazo</b>	<b>11</b>
<b>5. Posibles patrones de crecimiento</b>	<b>12</b>
5.1. Retorno de escala constante – Caso general . . . . .	17
<b>6. Retorno de escala constante – Caso matricial</b>	<b>21</b>
<b>7. Retornos de escala decrecientes</b>	<b>22</b>
7.1. Modelo de crecimiento de Solow . . . . .	25
7.2. Deducción algebraica de la función de producción de Cobb-Douglas . . . . .	26
7.3. Regresión lineal de la función de Cobb-Douglas para un caso en particular . . . . .	27
<b>8. El modelo de Solow sobre el crecimiento</b>	<b>28</b>
8.1. Estructura del mercado, dotaciones y limpieza del mercado . . . . .	28
<b>A. Seleccionar una medida de desempeño</b>	<b>30</b>
A.1. La ecuación normal . . . . .	31
A.2. Complejidad computacional . . . . .	32
A.3. Descenso del gradiente . . . . .	33

# 1. Introducción

Este proyecto hace referencia a la función de producción de Cobb-Douglas, siendo este publicado en [Douglas1976] 1928 por Douglas1976, quienes realizaron un estudio en el que se modeló el crecimiento de la economía estadounidense. Para este proyecto se desarrollará una aplicación con una base de datos como un caso particular, pero previo a su aplicación, se realizará una posible forma de cómo Charles Cobb y Paul Douglas llegaron a la formulación algebraica de la función. Al finalizar, se comparará la solución exacta de la ecuación con la obtenida por el método de mínimos cuadrados.

## 1.1. Función de producción de Cobb-Douglas

Para el análisis matemático de la función, es necesario describir los factores involucrados en el modelo.

## 1.2. Función de producción

Es la relación entre las cantidades máximas de productos que una empresa puede fabricar mediante el uso de diversas cantidades de insumos. Las múltiples funciones de producción están representadas por la combinación de factores de insumo (tecnología, capital, trabajo entre otros). Una función de producción se expresa como la ecuación (1) siguiente:

$$P = f(K, L, I) \quad (1)$$

donde:

- $P$  es la cantidad de producción.
- $K, L, I$  son los insumos.

# 2. El proyecto

En esta sección explicaré los detalles de mi proyecto y su visión. Espero que esta estructura se mejore considerablemente bajo la guía de mi mentor.

## 2.1. Cobb-Douglas y la función de producción ACMS

Partiendo de la función Cobb-Douglas

$$Y = bL^k C^{1-k}, \quad (2)$$

donde:

- $b$  representa el factor total de productividad,
- $Y$  la producción total,
- $L$  el trabajo,
- $C$  el capital.

Esta función fue generalizada siendo expresada de la manera siguiente:

$$f = \gamma x_1^{\alpha_1} \dots x_n^{\alpha_n}, \quad (3)$$

donde  $\gamma$  es una constante positiva y  $\alpha_1, \dots, \alpha_n$  son constantes no cero.

Se dice que una función de producción  $f$  es  $d$ -homogénea o homogénea de degradación  $d$ , si:

$$f(tx_1, \dots, tx_n) = t^d f(x_1, \dots, x_n), \quad (4)$$

Se mantiene para cada  $t \in \mathbb{R}$  en la función previamente definida.

Una función homogénea de degradación uno es llamado como linealmente homogéneo.

Si  $d > 1$ , la función homogénea mostrará un crecimiento a escala, caso contrario cuando  $d < 1$  mostrará un decrecimiento a escala.

Arrow, Chenery, Minhas y Solow(ACMS) introdujeron una función de producción de dos factores:

$$Q = F \cdot (aK^r + (1-a)L^r)^{1/r}, \quad (5)$$

donde:

- $Q$  representa el resultado,
- $F$  el factor de producción,
- $a$  el parámetro compartido,
- $k, L$  los factores de producción primario,
- $r = (s - 1) / s$ ,  $s = 1 / (1 - r)$  como la substitución de elasticidad.

La función de producción generalizada de ACMS se define:

$$f(x) = \gamma \sum_{i=1}^n (a_i^p x_i^p)^{d/p}, x = (x_1, \dots, x_n) \in D \subset \mathbb{R}_+^n, \quad (6)$$

con  $a_1, \dots, a_n, \gamma, p \neq 0$ , donde  $d$  es la degradación de homogeneidad.

Cabe resaltar que la función de producción homotética tiene la siguiente expresión como una función de producción:

$$f(x) = F(h(x_1, \dots, x_n)), \quad (7)$$

donde  $F$  es una función estrictamente creciente y  $h(x_1, \dots, x_n)$  es una función homogénea de cualquier degradación  $d$ . La función de producción homotética de la forma:

$$f(x) = \gamma \sum_{i=1}^n (a_i^p x_i^p)^{d/p}, \quad (\text{resp.}, \quad f(x) = F(x_1^{\alpha_1} \cdots x_n^{\alpha_n})), \quad (8)$$

es llamada función de producción ACMS generalizada homotética.

## 2.2. Breve descripción

Si  $f$  es una función de dos variables, a menudo dejamos que una letra como  $z$  denote el valor de  $f$  en el punto  $(x, y)$ , así  $z = f(x, y)$ . Entonces llamaremos a  $x$  e  $y$  las variables independientes, o los argumentos de  $f$ , mientras que  $z$  es llamada la variable independiente, porque el valor  $z$ , en general, depende de los valores  $x$  e  $y$ . El dominio de la función  $f$  es entonces el conjunto de todos los posibles pares de variables independientes, mientras que su rango es el conjunto de valores correspondientes de la variable dependientes. En economía,  $x$  e  $y$  son llamadas las variables exógenas, mientras que  $z$  es la variable endógena.

Una función de dos variables que aparecen en muchos modelos económicos es

$$F(x, y) = Ax^a y^b \quad (9)$$

donde  $A, a$  y  $b$  son constantes. Usualmente, uno asume que  $F$  es definida solo para  $x > 0$  e  $y > 0$ .

A función de la forma (9) es generalmente llamada la función de Cobb-Douglas. Se usa con mayor frecuencia para describir ciertos procesos de producción. Entonces  $x$  e  $y$  son llamados factores de entrada, mientras que  $F(x, y)$  es el número de unidades producidas, o la salida. En este caso  $F$  es llamada la función de producción.

Ejemplo 1 (Elasticidad de sustitución constante). Considere la “elasticidad de sustitución constante”, o la función CES

$$F(K, L) = A (aK^{-\rho} + (1 - a)L^{-\rho})^{-1/\rho} \quad (10)$$

donde  $A > 0$ ,  $K > 0$ ,  $L > 0$ ,  $a \in (0, 1)$ , y  $\rho \neq 0$ . Manteniendo  $A, K, L$  y  $a$  fijos, aplique la regla de L'Hôpital a  $z = \ln[F(K, L)/A]$  cuando  $\rho \rightarrow 0$  con el fin de mostrar que  $F(K, L)$  converge a la función de Cobb-Douglas  $AK^a L^{1-a}$ .

Solución. Obtenemos

$$z = \ln (aK^{-\rho} + (1 - a)L^{-\rho})^{-1/\rho} = -\ln (aK^{-\rho} + (1 - a)L^{-\rho}) / \rho \rightarrow \frac{0}{0} \text{ cuando } \rho \rightarrow 0.$$

Ya que  $(d/d\rho) K^{-\rho} = -K^{-\rho} \ln K$  y  $(d/d\rho) L^{-\rho} = -L^{-\rho} \ln L$ , aplicando la regla de L'Hôpital da

$$\begin{aligned} \lim_{\rho \rightarrow 0} z &= \lim_{\rho \rightarrow 0} \left[ \frac{aK^{-\rho} \ln K + (1 - a)L^{-\rho} \ln L}{aK^{-\rho} + (1 - a)L^{-\rho}} \right] \\ &= a \ln K + (1 - a) \ln L \\ &= \ln K^a L^{1-a}. \end{aligned}$$

Por lo tanto,  $e^z \rightarrow K^a L^{1-a}$ . De la definición de  $z$ , se sigue que  $F(K, L) \rightarrow AK^a L^{1-a}$  cuando  $\rho \rightarrow 0$ . □

Ejemplo 2 (Función de Cobb-Douglas). Una función de dos variables que aparece en muchos modelos económicos es

$$F(x, y) = Ax^a y^b \quad (11)$$

donde  $A$ ,  $a$  y  $b$  son constantes. Usualmente, uno asume que  $F$  está definida sola para  $x > 0$  e  $y > 0$ .

Una función  $F$  de la forma (11) es generalmente llamada la función de Cobb-Douglas<sup>1</sup>. Con frecuencia es usada para describir ciertos procesos productivos. Así,  $x$  e  $y$  son llamados los factores de entrada, mientras que  $F(x, y)$  es el número de unidades producidas, o la salida. En este caso,  $F$  es llamada una función de producción.

Ejemplo 3. Suponga que  $F(K, L)$  modela la producción de una empresa cuando sus entradas son capital y labor, respectivamente  $K$  y  $L$ . Una curva de nivel por esta función de producción es una curva en el plano  $KL$  dado por  $F(K, L) = Y_0$ , donde  $Y_0$  es una constante. Esta curva es llamada una isocuanta, que significa “igual cantidad”. Para una función de Cobb-Douglas,  $F(K, L) = AK^a L^b$ , con  $a + b < 1$  y  $A > 0$ , las figuras 3 y 3, respectivamente, muestra una parte de la gráfica cerca del origen, y tres de las isocuantas.

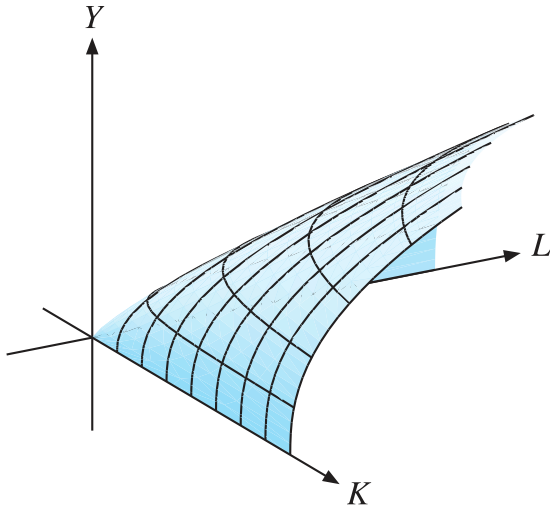


Figura 1: Gráfica de la función de producción Cobb-Douglas.

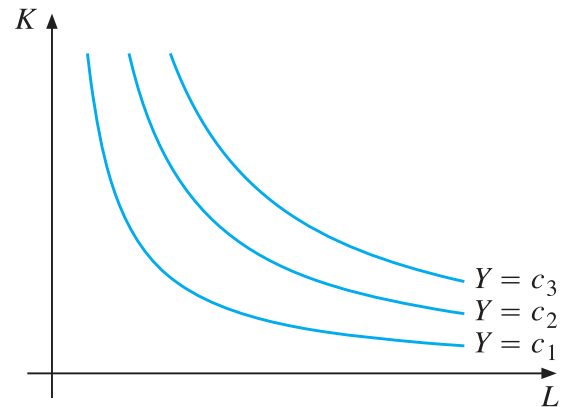


Figura 2: Isocuantas de la función de producción Cobb-Douglas.

Ejemplo 4 (Funciones  $n$ -lineales y log-lineales).

1. La demanda del azúcar en los Estados Unidos durante el período 1929–1936 fue estimado para ser descrito, aproximadamente, por la fórmula

$$x = 108.83 - 6.0294p + 0.164w - 0.4217t$$

donde  $x$  es la demanda del azúcar,  $p$  es su precio,  $w$  es un índice de producción y  $t$  es el año (donde  $t = 0$  corresponde a 1929).

2. La siguiente fórmula es una estimación para la demanda de cerveza en el Reino Unido:

$$x = 1.058x_1^{0.136}x_2^{-0.727}x_3^{0.914}x_4^{0.816}.$$

Aquí la cantidad demandada,  $x$ , es una función de cuatro variables:  $x_1$ , el ingreso per cápita,  $x_2$ , el precio de la cerveza,  $x_3$ , índice general de precios de productos básicos y  $x_4$ , la fuerza de la cerveza.

Las funciones más simples en el ejemplo anterior es la única en la parte (1). Las variables  $p$ ,  $w$  y  $t$  ocurren solo cuando a la primera potencia, y ellas son multiplicadas por constantes, no por cada otra. Tales funciones son llamadas lineales. En general

$$f(x_1, x_2, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n + b \quad (12)$$

<sup>1</sup>La función en (11) es llamada en honor a los investigadores americanos C.W.Cobb y P.H.Douglas, quien aplicaron esto, con  $a + b = 1$ , en un artículo científico que apareció en 1927 en la estimación de una función de producción. Sin embargo, debería correctamente ser llamada la “función de Wicksell”, porque el economista sueco K.Wicksell(1851-1926) introdujo tales funciones de producción antes de 1900.

donde  $a_1, a_2, \dots, a_n$  y  $b$  son constantes, es una función lineal en  $n$  variables.

La función en la parte (2) del ejemplo es un caso especial de la función general de Cobb-Douglas

$$F(x_1, x_2, \dots, x_n) = Ax_1^{a_1} x_2^{a_2} \dots x_n^{a_n} \quad (13)$$

donde  $A > 0$ ,  $a_1, \dots, a_n$  son constantes, definidas para  $x_1 > 0, x_2 > 0, \dots, x_n > 0$ . Note que al tomar el logaritmo natural a cada lado de (13) resulta

$$\ln F = \ln A + a_1 \ln x_1 + a_2 \ln x_2 + \dots + a_n \ln x_n. \quad (14)$$

Esto muestra que la función de Cobb-Douglas es log-lineal, ya que  $\ln F$  es una función lineal para  $\ln x_1, \ln x_2, \dots, \ln x_n$ .

Considere una función de producción Cobb-Douglas

$$Q = \alpha_0 \prod_{m=1}^M x_m^{\alpha_m}.$$

La maximización de ganancias con un precio de producción determinado de manera exógena exige que la empresa maximice la producción para un nivel de costo dado  $C$  (o minimizar los costos para una producción dada  $Q$ ). El lagrangeano para el problema de maximización es

$$\Lambda = \alpha_0 \prod_{m=1}^M x_m^{\alpha_m} + \lambda(C - \mathbf{p}'\mathbf{x}),$$

donde  $\mathbf{p}$  es el vector de los precios del factor  $M$ . Las condiciones necesarias para maximizar esta función son

$$\frac{\partial \Lambda}{\partial x_m} = \frac{\alpha_m Q}{x_m} - \lambda p_m = 0 \text{ y } \frac{\partial \Lambda}{\partial \lambda} = C - \mathbf{p}'\mathbf{x} = 0.$$

La solución conjunta proporciona  $x_m(Q, \mathbf{p})$  y  $\lambda(Q, \mathbf{p})$ . El costo total de producción es

$$\sum_{m=1}^M p_m x_m = \sum_{m=1}^M \frac{\alpha_m Q}{\lambda}.$$

El costo compartido asignado al  $m$ -ésimo factor es

$$\frac{p_m x_m}{\sum_{m=1}^M p_m x_m} = \frac{\alpha_m}{\sum_{m=1}^M \alpha_m} = \beta_m.$$

El modelo completo es<sup>2</sup>

$$\begin{aligned} \ln(C) &= \beta_0 + \beta_q \ln Q + \sum_{m=1}^M \beta_m \ln p_m + \varepsilon_c \\ s_m &= \beta_m + \varepsilon_m, \quad m = 1, \dots, M. \end{aligned}$$

Algebraicamente,  $\sum_{m=1}^M \beta_m = 1$  y  $\sum_{m=1}^M s_m = 1$ . (Este es el análisis de función de costo iniciado en el ejemplo 6.17. Regresaremos a esa aplicación a continuación.) Los costos compartidos también sumarán idénticamente a uno en los datos. Por lo tanto, se deduce que  $\sum_{m=1}^M \varepsilon_m = 0$  en cada dato punto por lo que el sistema es singular. Por el momento, ignore la función de costo. Deje que la  $M \times 1$  el vector de perturbación de las acciones será  $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M]'$ . Porque  $\varepsilon' \mathbf{i} = 0$ , donde  $\mathbf{i}$  es una columna de 1s, se deduce que  $E[\varepsilon \varepsilon' \mathbf{i}] = \sum \mathbf{i} = 0$ , lo que implica que  $\sum$  es singular.

Por lo tanto, los métodos de las secciones anteriores no se pueden usar aquí. (Debes verificar que la matriz de covarianza de muestra de los residuos de *OLS* también será singular).

La solución al problema de la singularidad parece ser soltar una de las ecuaciones, estimar el residuo y resolver el último parámetro del otro  $M - 1$ . La restricción  $\sum_{m=1}^M \beta_m = 1$  establece que la función de costo debe ser homogénea de grado uno en los precios. Si imponemos la restricción

$$\beta_M = 1 - \beta_1 - \beta_2 - \dots - \beta_{M-1}, \quad (15)$$

entonces el sistema se reduce a uno no singular,

$$\begin{aligned} \ln\left(\frac{C}{p_M}\right) &= \beta_0 + \beta_q \ln Q + \sum_{m=1}^{M-1} \beta_m \ln\left(\frac{p_m}{p_M}\right) + \varepsilon_c, \\ s_m &= \beta_m + \varepsilon_m, \quad m = 1, \dots, M - 1. \end{aligned}$$

Este sistema proporciona estimaciones de  $\beta_0, \beta_q$  y  $\beta_1, \dots, \beta_{M-1}$ . El último parámetro es estimado usando (15). Es irrelevante qué factor se elige como numerario; FGLS será invariante a qué factor se elige.

---

<sup>2</sup>Dejamos esto como ejercicio



Ejemplo 5 (Función de costo de Cobb-Douglas). El estudio de Nerlove (1963) sobre la industria de la energía eléctrica que examinamos en el ejemplo 6.6 proporciona una aplicación del modelo de función de costo Cobb-Douglas. Sus estimados de mínimos cuadrados ordinarios de los parámetros se enumeraron en el ejemplo 6.6. Entre los resultados hay (desafortunadamente) un negativo coeficiente de capital en tres de las seis regresiones. Nerlove también encontró que el simple modelo de Cobb-Douglas no era una cuenta adecuada para la relación entre el producto y el costo promedio. Christensen y Greene (1976) analizaron más a fondo los datos de Nerlove y argumentaron que el conjunto de datos con los datos de costos compartidos para estimar el sistema de demanda completo. Cuadro del apéndice F6.2 enumera las 145 observaciones de Nerlove con los datos de costos compartidos de Christensen y Greene. El costo es el costo total de generación en millones de dólares, la producción es en millones de kilovatios-hora, el precio del capital es un índice de costos de construcción, la tasa salarial es en dólares por hora para producción y mantenimiento, el precio del combustible es un índice del costo por BTU del combustible comprado por las empresas, y los datos reflejan los costos de producción de 1955. Las estimaciones de regresión se dan en la Tabla 10.4.

Los estimados de mínimos cuadrados de la función de costo Cobb-Douglas se dan en la primera columna. El coeficiente sobre el capital es negativo. Porque  $\beta_m = \beta_q \partial \ln Q / \partial \ln x_m$ , es decir, un positivo múltiplo de la elasticidad de salida del factor  $i$ -ésimo; este hallazgo es preocupante. La tercera columna presenta las restricciones de las estimaciones FGLS. Para obtener el estimador restringido, fijamos el modelo en forma del estimador SUR agrupado en (10-20),

$$y = \begin{bmatrix} \ln(C/P_f) \\ s_k \\ s_l \end{bmatrix} = \begin{bmatrix} i & \ln(Q) & \ln(P_k/P_f) & \ln(P_l/P_f) \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_q \\ \beta_k \\ \beta_l \end{pmatrix} + \begin{bmatrix} \epsilon_c \\ \epsilon_k \\ \epsilon_l \end{bmatrix}.$$

Note que esta formulación impone las restricciones  $\beta_1 = \alpha_3$  y  $\gamma_1 = \alpha_4$  en Existen 3 (145) = 435 observaciones en las matrices de datos. Las estimaciones de la economía de escala en el modelo de Cobb-Douglas son  $\frac{1}{\beta_q} = 1.39$  (primera columna) y (tercera columna), el cual sugiere algún incremento de retorno a escala. Nerlove, sin embargo, encontró evidencia que en tamaños de empresas extremadamente grandes, las economías a escalas disminuyeron y eventualmente desaparecieron. Para dar cuenta de esto (esencialmente una curva de costo promedio clásica en forma de U), él agregó un término cuadrático en la salida logarítmica en la función de costo. La única ecuación y los estimadores FGLS dieron en el segundo y cuarto conjunto de resultados.

La salida cuadrática da la función de costo promedio y la esperada forma de U. Podemos determinar el punto donde el costo promedio alcanza su mínimo al igualar  $\partial \ln C / \partial \ln Q$  a 1. Este es  $Q^* = \exp [(1 - \beta_q) / (2\beta_{qq})]$ . Usando los estimadores FGLS, su valor es  $Q^* = 4.669$ . Alrededor del 85% de las empresas en la muestra obtuvieron menos que eso, así por estos estimadores, muchas empresas en la muestra aún no habían agotado las economías de escala disponible. ??

Tenga en cuenta que esta formulación impone las restricciones  $\beta_1 = \alpha_3$  y  $\gamma_1 = \alpha_4$  en (10-4). Existen 3 (145) = 435 observaciones en las matrices de datos. El estimador es entonces FGLS, como se muestra en (10-22) Se agrega una columna adicional para el modelo de registro cuadrático. Dos cosas a tener en cuenta son los errores estándar dramáticamente más pequeños y la estimación ahora positiva (y razonable) de la coeficiente de capital Las estimaciones de las economías de escala en el modelo básico de Cobb-Douglas son  $1/\beta_q = 1.39$  (columna 1) y 1.31 (columna 3), lo que sugiere algunos rendimientos crecientes a escala. Sin embargo, Nerlove había encontrado evidencia de que a tamaños de empresa extremadamente grandes, economías de escala disminuyó y finalmente desapareció. Para dar cuenta de esto (esencialmente un clásico en forma de U-curva de costo promedio), agregó un término cuadrático en la producción de registro en la función de costo. los La ecuación única y las estimaciones de FGLS se dan en el segundo y cuarto conjuntos de resultados. El término de salida cuadrática le da a la función de costo promedio la forma de U esperada. Nosotros puede determinar el punto donde el costo promedio alcanza su mínimo al igualar  $d \ln C / d \ln Q$  a 1. Esto es  $Q^* = \exp [(1 - \beta_q) / (2\beta_{qq})]$ . Usando las estimaciones de FGLS, este valor es  $Q^* = 4,669$ . (La aplicación 5 considera usar el método delta para construir un intervalo de confianza para  $Q^*$ ). Alrededor de 85 de las empresas de la muestra tuvieron una producción inferior a esta, por lo que según estas estimaciones, la mayoría de las empresas de la muestra aún no habían agotado las economías de escala disponibles. La figura 3 muestra costos promedio predichos y reales para la muestra. (Para obtener una escala razonable, la más pequeña un tercio de las empresas se omiten de la figura.) Se calculan los costos promedio predichos en los promedios de muestra de los precios de entrada. La figura revela que eso más allá de un a pequeña escala, las economías de escala, aunque quizás estadísticamente significativas, son económicamente bastante pequeño.

		Mínimos cuadrados ordinarios		Restricciones factibles GLS	
Constante	$\beta_0$	-4.686 (0.885)	-3.764 (0.702)	-7.069 (0.107)	-5.707 (0.165)
ln Salida	$\beta_q$	0.721 (0.0174)	0.153 (0.0618)	0.766 (0.0154)	0.239 (0.0587)
ln <sup>2</sup> Salida	$\beta_{qq}$		0.0505 (0.0054)		0.0451 (0.00508)
ln $P_{\text{capital}}$	$\beta_k$	-0.0085 (0.191)	0.0739 (0.150)	0.424 (0.00946)	0.425 (0.00943)
ln $P_{\text{trabajo}}$	$\beta_l$	0.594 (0.205)	0.481 (0.161)	0.106 (0.00386)	0.106 (0.00380)
ln $P_{\text{combustible}}$	$\beta_f$	0.414 (0.0989)	0.445 (0.0777)	0.470 (0.0101)	0.470 (0.0100)

Cuadro 1: Estimados de la función de costos (Errores estimados estándar en paréntesis).

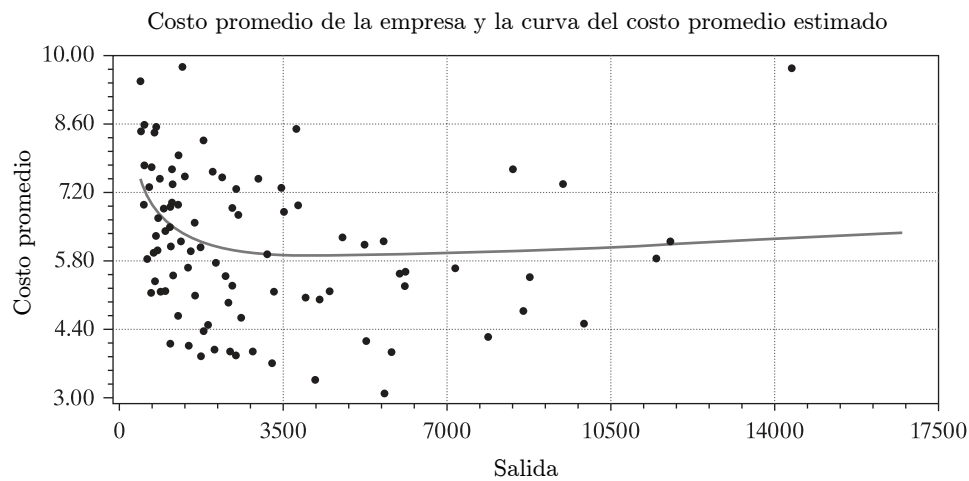


Figura 3: Gráfica del costo promedio predecido.

### 3. Introducción

Cualquier teoría depende de supuestos que no son del todo ciertos. Eso es lo que lo hace teoría. El arte de teorizar con éxito es hacer los supuestos simplificadores inevitables de tal manera que los resultados finales no sean muy sensibles. Una suposición “crucial” es una de las cuales las conclusiones dependen sensiblemente, y es importante que los supuestos cruciales sean razonablemente realistas. Cuando los resultados de una teoría parecen fluir específicamente de una suposición crucial especial, entonces, si la suposición es dudosa, los resultados son sospechosos.

Deseo argumentar que algo así es cierto en el modelo de crecimiento económico Harrod–Domar. La característica y poderosa conclusión de la línea de pensamiento Harrod–Domar es que incluso para el largo plazo, el sistema económico está en el mejor de los casos equilibrado sobre el filo del cuchillo del equilibrio del crecimiento. ¿Eran las magnitudes de los parámetros clave –la relación de ahorro, la relación capital-producto, la tasa de aumento de la mano de obra– si se deslizara un poco desde el punto muerto, la consecuencia sería un desempleo creciente o una inflación prolongada. En términos de Harrod, la cuestión crítica del equilibrio se reduce a una comparación entre la tasa natural de crecimiento que depende, en la ausencia del cambio tecnológico, en el aumento de la fuerza laboral, y la tasa de crecimiento garantizada que depende de los hábitos de ahorro e inversión de los hogares y las empresas.

Pero esta oposición fundamental de tasas garantizadas y naturales al final resulta que parte del supuesto crucial de que la producción tiene lugar en condiciones de proporciones fijas. No hay posibilidad de sustituir mano de obra por capital en producción. Si esta suposición se abandona, la noción del filo de cuchillo de equilibrio inestable parece ir con eso. De hecho, no es sorprendente que una rigidez tan grave en una parte del sistema implique falta de flexibilidad en otro.

Una característica notable del modelo Harrod–Domar es que estudia constantemente los problemas a largo plazo con las herramientas de corto plazo habitual. Normalmente se piensa en el largo plazo como el dominio del análisis neoclásico, la tierra del margen. En cambio Harrod y Domar hablan del largo plazo en términos del multiplicador, el acelerador, “el” coeficiente de capital. La mayor parte de este documento está dedicado a un modelo de crecimiento a largo plazo que acepta todos los supuestos de Harrod–Domar excepto el de proporciones fijas. En cambio supongo que la única mercancía compuesta es producida por trabajo y capital bajo las condiciones neoclásicas estándar. La adaptación del sistema a una tasa de incremento de la fuerza laboral dada de manera exógena se calcula en algún detalle, para ver si aparece la inestabilidad de Harrod. Las reacciones de interés precio-salario juegan un papel importante en este proceso de ajuste neoclásico, por lo que también se analizan. Luego, algunos de los otros rígidos supuestos se relajan ligeramente para ver qué cambios cualitativos resultan: se permite un cambio tecnológico neutral y un interés elástico horario de ahorro. Finalmente, las consecuencias de ciertas relaciones y rigideces más “keynesianas” son brevemente consideran.

### 4. Un modelo de crecimiento a largo plazo

Solo hay una mercancía, la producción como un todo, cuya tasa de producción se designa  $Y(t)$ . Así podemos hablar inequívocamente del ingreso real de la comunidad. Parte de cada salida instantánea es consumida y el resto se ahorra e invierte. La fracción de la salida ahorrada es una constante  $s$ , de modo que la tasa de ahorro es  $sY(t)$ . El stock de capital de la comunidad  $K(t)$  toma la forma de una acumulación de la mercancía compuesta. La inversión neta es solo la tasa de aumento de este capital social  $dK/dt$  o  $\dot{K}$ , por lo que tenemos la identidad básica en cada instante de tiempo:

$$\dot{K} = sY \quad (16)$$

La salida es producida con la ayuda de dos factores de producción, capital y trabajo, cuya tasa de ingreso es  $L(t)$ . Las posibilidades tecnológicas son representadas por una función de producción.

$$Y = F(K, L) \quad (17)$$

La salida es entendida como la salida neta después de hacer buena la depreciación del capital. Sobre la producción, todo lo que diremos en este momento es que muestra rendimientos constantes a escala. Por lo tanto, la función de producción es homogénea de primer grado. Esto equivale a asumir que no existe un recurso escaso no aumentable como la tierra. Retornos de escala constante parece la suposición natural para hacer en una teoría de crecimiento. El caso de tierras escasas conduciría a rendimientos decrecientes a escala en capital y trabajo y el modelo se volvería más Ricardiano.

Insertando (17) en (16) obtenemos

$$\dot{K} = sF(K, L). \quad (18)$$

Este es una ecuación con dos incógnitas. Una primera manera de acercarse al sistema sería agregar una ecuación de demanda de trabajo: la productividad física del trabajo marginal es igual a la tasa salarial real; y una ecuación de

oferta de trabajo. Este último podría tomar la forma general de hacer trabajo proporcionar una función del salario real, o más clásico de poner el salario real igual a un nivel de subsistencia convencional. En cualquier caso serían tres ecuaciones en las tres incógnitas  $K$ ,  $L$ , salario real.

En cambio, procedemos más en el espíritu del modelo Harrod. Como un resultado exógeno del crecimiento de la población, la fuerza laboral aumenta a una tasa relativa constante  $n$ . En ausencia de cambio tecnológico,  $n$  es la tasa natural de crecimiento de Harrod. Así:

$$L(t) = L_0 e^{nt} \quad (19)$$

En (18)  $L$  representa el empleo total; en (19)  $L$  representa la oferta de trabajo disponible. Al identificar los dos estamos asumiendo que el empleo se mantiene perpetuamente. Cuando insertamos (19) en (18) obtenemos

$$\dot{K} = sF(K, L_0 e^{nt}) \quad (20)$$

tenemos la ecuación básica que determina el camino temporal de la acumulación del capital que debe ser seguida si todas los trabajos disponibles están empleados.

Alternativamente, (19) puede ser visto como una curva de oferta de mano de obra. Eso dice que la fuerza laboral que crece exponencialmente se ofrece para un empleo completamente inelástico. La curva de oferta de trabajo es una línea vertical que se mueve hacia la derecha en el tiempo a medida que la fuerza laboral crece de acuerdo para (19). Luego, la tasa salarial real se ajusta para que toda la mano de obra disponible sea empleado, y la ecuación de productividad marginal determine la tasa salarial que realmente gobernará.

En resumen, (20) es una ecuación diferencial con la única variable  $K(t)$ . Su solución da el único perfil de tiempo del capital social de la comunidad que empleará plenamente la mano de obra disponible. Una vez que nosotros conozca el camino temporal del stock de capital y el de la fuerza laboral, podemos calcular desde la función de producción la ruta de tiempo correspondiente de salida real. La ecuación de productividad marginal determina la trayectoria temporal del salario real. También hay una suposición involucrada de pleno empleo del stock de capital disponible. En cualquier punto de tiempo en que el stock de capital preexistente (el resultado de una acumulación previa) se suministra de manera inelástica. Por lo tanto, existe una ecuación de productividad marginal similar para el capital que determina el alquiler real por unidad de tiempo para los servicios de capital social. El proceso puede ser visto de esta manera: en cualquier momento la oferta laboral disponible está dado por (19) y el stock de capital disponible también es un dato. Ya que el rendimiento real de los factores se ajustará para lograr el pleno empleo de trabajo y capital podemos usar la función de producción (17) para encontrar la tasa actual de salida. Entonces la propensión a ahorrar nos dice cuánto de la producción neta se ahorrará e invertirá. Por eso conocemos la acumulación del capital neta durante el período actual. Agregado al stock ya acumulado, esto da el capital disponible para el próximo período, y todo el proceso puede repetirse.

## 5. Posibles patrones de crecimiento

Para ver si siempre existe una ruta de acumulación de capital consistente con cualquier tasa de crecimiento de la fuerza laboral, debemos estudiar la ecuación diferencial (20) por la naturaleza cualitativa de sus soluciones. Naturalmente sin especificar la forma exacta de la función de producción no podemos esperar encontrar la solución exacta. Pero ciertas propiedades amplias son sorprendentemente fáciles de aislar, incluso gráficamente.

Para ello, introducimos una nueva variable  $r = \frac{K}{L}$ , la relación de capital al trabajo. Por lo tanto, tenemos  $K = rL = rL_0 e^{nt}$ . Diferenciando con respecto al tiempo que tenemos

$$\dot{K} = L_0 e^{nt} r' + nrL_0 e^{nt}. \quad (21)$$

Reemplazando esto en (20):

$$(\dot{r} + nr)L_0 e^{nt} = sF(K, L_0 e^{nt}).$$

Pero debido al retorno de escala constante podemos dividir ambas variables en  $F$  por  $L = L_0 e^{nt}$ , no obstante, multiplicamos  $F$  por el mismo factor. Así

$$(\dot{r} + nr)L_0 e^{nt} = sL_0 e^{nt} F\left(\frac{K}{L_0 e^{nt}}, 1\right)$$

y dividiendo el factor común llegamos finalmente a

$$\dot{r} = sF(r, 1) - nr. \quad (22)$$

Aquí tenemos una ecuación diferencial que involucra solamente la relación capital-trabajo.

Esta ecuación fundamental se puede alcanzar menos formalmente. Como  $r = \frac{K}{L}$ , la tasa de cambio relativa de  $r$  es la diferencia entre las tasas relativas de cambio de  $K$  y  $L$ . Eso es:

$$\frac{\dot{r}}{r} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L}.$$

Ahora primero que nada  $\frac{\dot{L}}{L} = n$ . En segundo lugar,  $\dot{K} = sF(K, L)$ . Haciendo estas substituciones:

$$\dot{r} = r \frac{sF(K, L)}{K} - nr.$$

Ahora divida  $L$  de  $F$  como antes, note que que  $\frac{L}{K} = \frac{1}{r}$  y obtenemos (22) nuevamente.

La función  $F(r, 1)$  que aparece en (22) es fácil de interpretar. Esta es la curva del producto total cuando varían las cantidades  $r$  de capital con una unidad de trabajo. Alternativamente, da salida por trabajador como una función de capital por trabajador. Así (22) establece que la tasa del cambio de la relación capital-trabajo es la diferencia de dos términos, uno representando el incremento de capital y uno el incremento de trabajo.

Cuando  $\dot{r} = 0$ , la relación capital-trabajo es una constante, y el capital existente debe expandirse al mismo ritmo que la fuerza laboral, es decir,  $n$ .

(La tasa de crecimiento garantizada, garantizada por la tasa real apropiada de retorno al capital, es igual a la tasa natural.) En la Figura I, el rayo que pasa por el origen con pendiente  $n$  representa la función  $nr$ . La otra curva es la función  $sF(r, 1)$ . Aquí se dibuja para pasar por el origen y convexo hacia arriba: sin salida a menos que ambas entradas sean positivas, y la disminución de la productividad marginal del capital, como sería el caso, por ejemplo, con la función Cobb-Douglas. En el punto de intersección  $nr = sF(r, 1)$  y  $\dot{r} = 0$ . Si la relación capital-trabajo  $r^*$  debe establecerse, se mantendrá, y el capital y el trabajo crecerá de allí en adelante en proporción. Por la constante retornos a escala

Formalmente, una función de producción se define para tener:

- Constante retorno a escala si (para cualquier constante  $a$  es mayor que 0)  $F(aK, aL) = aF(K, L)$  (Función  $F$  es homogénea de grado 1).
- Retornos a escala crecientes si (para cualquier constante mayor que 1)  $F(aK, aL) > aF(K, L)$ .
- Retornos a escala decrecientes si (para cualquier constante  $a$  mayor que 1)  $F(aK, aL) < aF(K, L)$ .

donde  $K$  y  $L$  son factores de producción-capital y trabajo, respectivamente.

En una configuración más general, para procesos de producción de múltiples entradas y múltiples salidas, se puede suponer que la tecnología se puede representar a través de algún conjunto de tecnología, llámelo  $T$  que debe satisfacer algunas condiciones de regularidad de la teoría de la producción. En este caso, la propiedad de retorno de escala constante es equivalente a decir que el conjunto tecnológico es un cono, es decir, satisface la propiedad  $aT = T$ ,  $\forall a > 0$ . A su vez, si hay una función de producción que describirá el conjunto de tecnología  $T$ , deberá ser homogéneo de grado 1.

**Definición 1 (Rendimiento de escala).** La forma funcional de Cobb-Douglas tiene una constante retorno de escala cuando la suma de sus exponentes es 1. En este caso, la función es

$$F(K, L) = AK^b L^{1-b} \quad (23)$$

donde  $A > 0$  y  $0 < b < 1$ . Así

$$F(aK, aL) = A(aK)^b (aL)^{1-b} = Aa^b a^{1-b} K^b L^{1-b} = aAK^b L^{1-b} = aF(K, L).$$

Aquí como entrada usamos todas las escalas por un factor multiplicador  $a$ , la salida también escala por  $a$  y así existen constantes de retorno de escala.

Pero, si la función de producción de Cobb-Douglas tiene su forma general

$$F(K, L) = AK^b L^c \quad (24)$$

donde  $0 < b < 1$  y  $0 < c < 1$ , entonces existen retornos crecientes si  $b + c > 1$ , pero retornos decrecientes si  $b + c < 1$ , dado que

$$F(aK, aL) = A(aK)^b (aL)^c = Aa^b a^c K^b L^c = a^{b+c} AK^b L^c = a^{b+c} F(K, L),$$

que para  $a > 1$  es mayor que o menor que  $aF(K, L)$  cuando  $b + c$  es mayor o menor que uno.

Hay dos clases especiales de funciones de producción que a menudo se analizan. La función de producción  $Q = f(X_1, X_2, \dots, X_n)$  se dice que es homogéneo de grado  $m$ , si se le da alguna constante positiva  $k$ ,  $f(kX_1, kX_2, \dots, kX_n) = k^m f(X_1, X_2, \dots, X_n)$ . Si  $m > 1$ , la función exhibe rendimientos crecientes a escala, y exhibe rendimientos decrecientes a escala si  $m < 1$ . Si es homogéneo de grado 1, exhibe rendimientos constantes a escala. La presencia de rendimientos crecientes significa que un aumento del uno por ciento en los niveles de uso de todas las entradas daría como resultado un aumento de más del uno por ciento en la producción; la presencia de rendimientos decrecientes significa que daría como resultado un aumento de producción de menos del uno por ciento. Los retornos constantes a escala son el caso intermedio. En la función de producción Cobb-Douglas mencionada anteriormente, los rendimientos a escala aumentan si  $a_1 + a_2 + \dots + a_n > 1$ , disminuyendo si  $a_1 + a_2 + \dots + a_n < 1$ , y constante si  $a_1 + a_2 + \dots + a_n = 1$ .

Si una función de producción es homogénea y de grado uno, este a veces llamada “linealmente homogénea”. Una función de producción linealmente homogénea con entradas capital y labor tienen las propiedades de que los productos físicos marginales y promedio tanto del capital como del trabajo pueden expresarse solamente como funciones de la relación capital-trabajo. Además, en este caso, si cada entrada se paga a una tasa igual a su producto marginal, los ingresos de la empresa se agotarán exactamente y no habrá ganancias económicas excesivas.

Las funciones homotéticas son funciones cuya tasa de sustitución técnica marginal (la pendiente de la isocuanta, una curva dibujada a través del conjunto de puntos en dicho espacio de trabajo-capital en el que se produce la misma cantidad de producción para combinaciones variables de las entradas) es homogénea de grado cero. Debido a esto, a lo largo de los rayos que provienen del origen, las pendientes de las isocuantas serán las mismas. Las funciones homotéticas tienen la forma  $F(h(X_1, X_2))$  donde  $F(y)$  es una función monótona creciente (la derivada de  $F(y)$  es positiva  $dF/dy > 0$ , y la función  $h(X_1, X_2)$  es una función homogénea de cualquier grado.

La elasticidad de sustitución constante (CES), en economía, es una propiedad de algunas funciones de producción y funciones de utilidad.

Específicamente, este es un tipo particular de función agregado que combina dos o más tipos de productos de consumos, o dos o más tipos de entradas de producción dentro de un cantidad agregado. Esta función de agregación exhibe una elasticidad de sustitución constante.

Definición 2 (Elasticidad de sustitución constante). La función de producción CES es una función de producción neoclásica que muestra una elasticidad de sustitución constante. En otras palabras, la producción tecnológica tiene un porcentaje de cambio constante en factores (por ejemplo, trabajo y capital) proporcional debido al cambio porcentual en la tasa marginal de la sustitución técnica. Los dos factores (capital y trabajo) de la función de producción fue introducido por Solow y más tarde popularizado por Arrow, Chenery, Minhas y Solow es

$$Q = F \cdot (a \cdot K^\rho + (1 - a) \cdot L^\rho)^{\frac{v}{\rho}} \quad (25)$$

donde

- $Q$  es la cantidad de salida,
- $F$  es el factor de productividad,
- $a$  es el parámetro forma,
- $K, L$  son las cantidades de los factores de producción primario (capital y trabajo)
- $\rho = \frac{\sigma-1}{\sigma}$  es el parámetro de sustitución,
- $\sigma = \frac{1}{1-\rho}$  es elasticidad de sustitución,
- $v$  es el grado de homogeneidad de la función de producción. Donde  $v = 1$  es el retorno de escala constante,  $v < 1$  es el retorno de escala decreciente y  $v > 1$  es el retorno de escala creciente.

Como su nombre lo sugiere, la función de producción CES exhibe una elasticidad de sustitución constante entre el capital y el trabajo. Leontief, linear y las funciones de Cobb-Douglas son casos especiales de la función de producción CES. Esto es,

- Si  $\rho$  se aproxima a 1, tenemos una lineal o función de sustituto perfecto.
- Si  $\rho$  se aproxima a cero en el límite, obtenemos la función de producción de Cobb-Douglas.
- Si  $\rho$  se aproxima al menos infinito, obtenemos la Leontief o función de producción perfecta complementaria.

La forma general de la función de producción CES, con  $n$  entradas, es

$$Q = F \cdot \left[ \sum_{i=1}^n a_i X_i^r \right]^{\frac{1}{r}} \quad (26)$$

donde

- $Q$  es cantidad de salida
- $F$  es el factor de productividad
- $a_i$  es el parámetro forma de la entrada  $i$ ,  $\sum_{i=1}^n a_i = 1$
- $X_i$  son las cantidades de los factores de producción,  $i = 1, 2, \dots, n$ .
- $s = \frac{1}{1-r}$  es la elasticidad de sustitución.

Extendiendo la forma función CES (Solow) para acomodar los múltiples factores de producción crea algunos problemas. Sin embargo, no existe una forma completamente general para hacer esto. Uzawa mostró que solo  $n$  factores posibles de la función de producción  $n > 2$  con elasticidades de sustitución parciales constantes requiere o todas las elasticidades entre pares de factores son idénticas, o si alguna difiere, todos ellos deben ser igual a cada otra y todas las elasticidades restantes deben ser unitarias. Esto es verdad para cualquier función de producción. Esto significa el uso de la forma funcional CES para más dos factores significará general que no existe una elasticidad de sustitución entre todos los factores.

Las funciones CES anidadas son comúnmente encontradas en los modelos de equilibrio parcial y equilibrio general. Diferentes anidamientos (niveles) permiten la introducción de las elasticidades de sustitución apropiadas.

Definición 3 (Función de utilidad CES). La misma forma funcional CES alcanza como una función de utilidad en la teoría del consumidor. Por ejemplo, si existen  $n$  tipos de productos de consumos  $x_i$ , entonces el consumo agregado  $X$  podría definirse usando el agregado CES:

$$X = \left[ \sum_{i=1}^n a_i^{\frac{1}{s}} x_i^{\frac{s}{s-1}} \right]^{\frac{s-1}{s}} \quad (27)$$

Aquí nuevamente, los coeficientes  $a_i$  son los parámetros forma y  $s$  es la elasticidad de sustitución. Por lo tanto, los productos de consumo  $x_i$  son perfectos sustitutos cuando  $s$  se aproxima al infinito y complemento perfecto cuando  $s$  se aproxima a cero. El agregado CES es también algunas veces llamado el agregador Armington, el cual fue discutido por Armington (1969).

Las funciones de utilidad CES son un caso especial de las preferencias homotéticas.

El siguiente es un ejemplo de la función de utilidad CES para dos productos,  $x$  e  $y$  con igualdad compartida:

$$u(x, y) = (x^r + y^r)^{1/r}. \quad (28)$$

La función expendidora en el caso es:

$$e(p_x, p_y, u) = \left( p_x^{r/(r-1)} + p_y^{r/(r-1)} \right)^{(r-1)/r} \cdot u. \quad (29)$$

La función de utilidad indirecta tiene su inversa:

$$v(p_x, p_y, I) = \left( p_x^{r/(r-1)} + p_y^{r/(r-1)} \right)^{(1-r)/r} \cdot I. \quad (30)$$

La funciones de demanda son:

$$x(p_x, p_y, I) = \frac{p_x^{1/(r-1)}}{p_x^{r/(r-1)} + p_y^{r/(r-1)}} \cdot I$$

$$y(p_x, p_y, I) = \frac{p_y^{1/(r-1)}}{p_x^{r/(r-1)} + p_y^{r/(r-1)}} \cdot I$$

La función de utilidad CES es uno de los casos considerados por Dixit y Stiglitz (1977) en su estudio de la diversidad del producto optimal en el contexto de la competición monopolística.

Note que la diferencial entre la utilidad CES y la utilidad isoelástica: La función de utilidad CES es una función de utilidad ordinal que representa las preferencias sobre consumo seguro mientras que la función de utilidad isoelástica es una función de utilidad cardinal que representa en loterías. Una función de utilidad CES indirecta (dual) ha sido usado para derivar la marca de consistencia-utilidad de sistemas donde la demanda categórica son determinadas endógenamente por un multicategorizador, la función de utilidad CES indirecto. Esto también se ha mostrado que las preferencias son autoduales y ambos son primales y duales podrían exhibir cualquier grado de convexidad.



La existencia y la estabilidad relativa de un único crecimiento balanceado para modelos multisectoriales fueron establecidos por Solow y Samuelson bajo el supuesto de retorno de escala constante. Ellos estudiaron dos tipos de sistemas de ecuaciones: el sistema de ecuación en diferencias y el sistema de ecuación diferencial. Later Muth y Suit estudiaron el sistema formado bajo el supuesto de retorno de escala decreciente. El primer objetivo de este artículo es estudiar algún sistema de ecuación diferencial bajo los supuestos más débiles que los impuestos por Solow y Samuelson, pero que retenga el supuesto de retorno constante de escala. El segundo objetivo es investigar cierto sistema de ecuación diferencial bajo el supuesto de retorno de escala decreciente.

## 5.1. Retorno de escala constante – Caso general

Nuestro sistema es expresado por las siguientes ecuaciones:

$$\dot{X}_i = H^i(X_1, \dots, X_n), \quad (i = 1, \dots, n). \quad (31)$$

El sistema de arriba es modelo de crecimiento balanceado de Solow–Samuelson. Los  $H^i$ 's son definidos para cualquier  $(X_1, \dots, X_n) \geq 0$  y son asumidos que son continuas con respecto a cualquier variable y positivamente homogénea de grado uno. A lo largo del artículo, los  $X_i$ 's son restringidos a valores no negativos. Además, las funciones son solo definidas para valores no negativos. Esto es asumido que

$$H^i \text{ es no decreciente en todas las variables, excepto en } X^i, \quad (32)$$

y que

$$\{H^1, \dots, H^n\} \text{ es indescomponible.} \quad (33)$$

Aquí la indescomposibilidad es definido como en Morishima. Esto es, para cualquier conjunto de índices  $R = \{i_1, \dots, i_r\}$ , las relaciones  $X_i = X'_i$  para  $i \in R$  y  $X_l < X'_l$  para  $l \notin R$  implica que existe por lo menos un  $i \in R$  tal que  $H^i(X_1, \dots, X_n) < H^i(X'_1, \dots, X'_n)$ . Requerimos que  $H^i$  sea no decreciente en  $X_j$ , para  $j \neq i$ , sin la restricción sobre la dependencia de  $H^i$  sobre  $X_i$ . En contraste del supuesto de Solow y Samuelson que  $H^i$  es creciente en todos los  $X_j$ .

Dadas sus supuestos y la homogeneidad de  $H^i$  ( $i = 1, \dots, n$ ), este sigue que  $H^i \geq 0$  ( $i = 1, \dots, n$ ) para  $X_j \geq 0$  ( $j = 1, \dots, n$ ), y que,  $H^i = 0$  para todo  $i$ , si y solo si  $X_j = 0$  para todo  $j$ . En nuestro caso, sin embargo,  $H^i$  no es necesariamente creciente en  $X$ . Por ello, no podemos obtener las propiedades mencionadas arriba. Así, asumimos ellos. Esto es, podemos asumir que

$$H^i \geq 0 \quad (i = 1, \dots, n) \text{ para } X_j \geq 0 \quad (j = 1, \dots, n). \quad (34)$$

Entonces, de la indescomposibilidad y la homogeneidad de  $H^i$ ,  $H^i = 0$  para todo  $i$ , si y solo si  $X_j = 0$  para todo  $j$ . Nuestro ánimo es probar el siguiente teorema.

**Teorema 1.** Para el sistema de ecuaciones diferenciales, existe un único determinado positivo autovalor, estrictamente un único positivo autovector normalizado y así un único camino de crecimiento balanceado. Más aún, cualquier solución del camino del sistema relativamente se aproxima al camino de crecimiento balanceado.

**Demostración.** Podemos mostrar por un procedimiento similar al de Solow y Samuelson sobre la existencia de un autovalor positivo  $\lambda$  y un autovector no negativo, no nulo  $V = (V_1, V_2, \dots, V_n)$  tal que

$$\begin{aligned} \lambda V_1 &= H^1(V_1, \dots, V_n), \\ &= \vdots \\ \lambda V_n &= H^n(V_1, \dots, V_n). \end{aligned}$$

Mostraremos que todas las componentes del autovector  $V$  son positivas. Suponga que algunas componentes de  $V$  son ceros. Sin pérdida de generalidad, podríamos suponer que

$$V_i = 0 \quad \text{para } i \leq r (< n),$$

y

$$V_i > 0 \quad \text{para } n \geq i > r.$$

Entonces,

$$\begin{aligned}
0 &= H^1(0, \dots, 0, V_{r+1}, \dots, V_n), \\
&= \vdots 0 &= H^r(0, \dots, 0, V_{r+1}, \dots, V_n), \\
0 < \lambda V_{r+1} &= H^{r+1}(0, \dots, 0, V_{r+1}, \dots, V_n), \\
&= \vdots 0 < \lambda V_n &= H^n(0, \dots, 0, V_{r+1}, \dots, V_n).
\end{aligned}$$

Pero esto contradice la suposición de indescomposabilidad, así es fácilmente visto haciendo

$$\begin{aligned}
R &\equiv \{1, \dots, r\}, \\
(X_1, \dots, X_r, X_{r+1}, \dots, X_n) &\equiv (0, \dots, 0, V_{r+1}, \dots, V_n), \\
(X'_1, \dots, X'_r, X'_{r+1}, \dots, X'_n) &\equiv (0, \dots, 0, 2V_{r+1}, \dots, 2V_n).
\end{aligned}$$

Ahora, mostraremos la unicidad del autovalor. Suponga que existe otra tupla de un autor valor positivo y un autovector  $(\mu, U)$ . Entonces obtenemos los siguientes conjuntos de relaciones

$$\lambda = H^1\left(1, \frac{V_2}{V_1}, \dots, \frac{V_n}{V_1}\right), \quad (35)$$

$$\lambda = H^2\left(\frac{V_1}{V_2}, 1, \dots, \frac{V_n}{V_2}\right), \quad (36)$$

$$= \vdots \quad (37)$$

$$\lambda = H^n\left(\frac{V_1}{V_n}, \frac{V_2}{V_n}, \dots, 1\right), \quad (38)$$

$$\mu = H^1\left(1, \frac{U_2}{U_1}, \dots, \frac{U_n}{U_1}\right), \quad (39)$$

$$\mu = H^2\left(\frac{U_1}{U_2}, \frac{U_2}{U_n}, \dots, 1\right). \quad (40)$$

Asuma que  $\lambda > \mu$ . Compare Entonces,

$$H^1\left(1, \frac{V_2}{V_1}, \dots, \frac{V_n}{V_1}\right) > H^1\left(1, \frac{U_2}{U_1}, \dots, \frac{U_n}{U_1}\right).$$

Dado que  $H^1$  es no decreciente en todos los argumentos, excepto en el primero, podemos reemplazar  $i = 2$ . Esto es,

$$\frac{V_2}{V_1} > \frac{U_2}{U_1}. \quad (41)$$

Compare Entonces,

$$H^2\left(\frac{V_1}{V_2}, 1, \dots, \frac{V_n}{V_2}\right) > H^2\left(\frac{U_1}{U_2}, 1, \dots, \frac{U_n}{U_2}\right).$$

Dado que  $V_1/V_2 < U_1/U_2$ , y  $H^2$  es no decreciente en todos los argumentos, excepto en el segundo, debemos tener, digamos,

$$\frac{V_3}{V_2} > \frac{U_3}{U_2}. \quad (42)$$

De obtenemos  $V_1/V_3 < U_1/U_3$  y  $V_2/V_3 < U_2/U_3$ . Continuando con este razonamiento, alcanzamos una contradicción para las últimas relaciones

Dado que los argumentos diagonales en el lado de derecho de ambos grupos de relaciones son todos uno, no necesitamos asumir que  $H^i$  es creciente en  $X^i$ . El razonamiento de arriba ha sido alcanzado usado por Solow y Samuelson para mostrar la unicidad de los autovalores para el caso  $n = 2$ . Pero ellos usan diferentes razonamientos para el caso general. En este razonamiento, ellos usan la propiedad que  $H^i$  es creciente en  $X_j$ .

Notamos también que el razonamiento de arriba es usado por Solow y Samuelson para mostrar la unicidad del vector normalizado y que el procedimiento es aplicable con un ligera modificación en nuestro caso también. Así, podemos omitir la prueba de  $V = \alpha U$ . Aquí,  $\alpha$  es una constante de proporcionalidad.

Nuestro siguiente objetivo es mostrar que la estabilidad relativa del camino dinámico. Definimos nuevas variables,

$$y_i = \frac{X_i}{V_i e^{\lambda t}}, \quad (i = 1, \dots, n). \quad (43)$$

Entonces,

$$y_i V_i e^{\lambda t} = X_i.$$

Diferenciando ambos lados de esta relación, obtenemos

$$\dot{y}_i V_i e^{\lambda t} + \lambda y_i V_i e^{\lambda t} = \dot{X}_i \quad (i = 1, \dots, n). \quad (44)$$

Sustituyendo las relaciones dentro del sistema original, obtenemos

$$\dot{y}_i = H^i \left( \frac{V_1}{V_i} y_1, \dots, \frac{V_n}{V_i} y_n \right) - \lambda y_i, \quad (i = 1, \dots, n). \quad (45)$$

Ponga

$$\min \{y_i(t)\} = m(t) = y_{k_1}(t) = \dots = y_{k_r}(t),$$

y suponga que

$$y_\ell(t) > m(t) \quad \text{para } \ell \neq k_j.$$

Entonces,

$$\dot{y}_{k_j}(t) \geq 0 \quad \text{para todo } j \leq r$$

y

$$\dot{y}_{k_j}(t) > 0 \quad \text{para al menos un } j \leq r.$$

Esto es mostrado como sigue.

$$\begin{aligned} \dot{y}_{k_j} &= H^{k_j} \left( \frac{V_1}{V_{k_j}} y_1, \dots, \frac{V_n}{V_{k_j}} y_n \right) - \lambda y_{k_j} \\ &\geq H^{k_j} \left( \frac{V_1}{V_{k_j}} m(t), \dots, \frac{V_n}{V_{k_j}} m(t) \right) - \lambda m(t) \\ &= m(t) H^{k_j} \left( \frac{V_1}{V_{k_j}}, \dots, \frac{V_n}{V_{k_j}} \right) - \lambda m(t) = 0, \quad \text{para } j = 1, \dots, r. \end{aligned}$$

Pero la desigualdad se mantiene para al menos un  $k_j$ . Esto sigue de la suposición de indescomposibilidad si ponemos

$$\begin{aligned} R &\equiv \{k_1, \dots, k_r\} \\ (X_1, \dots, X_n) &= (V_1 m(t), \dots, V_n m(t)) \end{aligned}$$

y

$$(X'_1, \dots, X'_n) = (V_1 y_1, \dots, V_n y_n).$$

Con esta propiedad, inferimos que el mínimo valor de  $y_i(t)$  no puede mantenerse constante por siempre. Para, cada momento de tiempo, el número de mínimos  $y_k(t)$ 0s es decreciente. Eventualmente, existe solo un mínimo  $y_k(t)$ . Por ello, el mismo mínimo debe incrementar. Dado que el lapso de tiempo continuamente en nuestro caso,  $m(t)$  siempre incrementa sobre el tiempo, provisto que  $y_\ell(t) > m(t)$  para al menos un  $\ell$ . Esto es posible que

$$\frac{dm(t)}{dt} = 0,$$

en un cierto punto. Pero  $m(t)$  se mantiene constante solo por un corto periodo infinitesimal. Eso no hace el residuo estacionario para un periodo finito. La figura 1 muestra la situación. Ponga

$$\max_i \{y_i(t)\} = M(t).$$

Entonces, podemos mostrar que  $M(t)$  decrece, provisto por  $Y_\ell(t) < M(t)$  para al menos un  $\ell$ .

Así,  $m(t)$  incrementa y converge a un cierto valor positivo  $m^*$  y  $M(t)$  decrece y converge a cierto valor positivo  $M^*$ . Esto es,

$$\lim_{t \rightarrow \infty} m(t) = m^*.$$

$$\lim_{t \rightarrow \infty} M(t) = M^*.$$

Entonces,

$$m^* \leq M^*.$$

Tenemos que probar que

$$m^* = M^*.$$

Suponga que  $m^* < M^*$ . Considere un conjunto de vectores en el espacio  $n$ -dimensional que

$$S \equiv \{y \equiv (y_1, \dots, y_n) : \min_i y_i = m^* \text{ y } \max_i y_i = M^*\}.$$

Este es un conjunto compacto. Considere un camino dinámico que empieza de un punto en este conjunto. Entonces, por el mismo razonamiento de arriba, el mínimo valor de los  $y_i$ 's incrementa y el máximo valor de los  $y_i$ 's decrece. Para hacer explícito esa dependencia en el valor inicial de  $y$  en  $S$ , escribimos, respectivamente,

$$m^*(t; y) \text{ y } M^*(t, y).$$

Luego,

$$m^*(\tau, y) > m^*(0, y) = m^* \text{ y } M^*(\tau, y) < M^*(0, y) = M^*.$$

Aquí,  $\tau$  es un valor positivo arbitrariamente escogido. Pero,

$$\inf_{y \in S} \{m^*(\tau, y) - m^*(0, y)\} = \varepsilon$$

y

$$\inf_{y \in S} \{M^*(0, y) - M^*(\tau, y)\} = \delta.$$

Dado que  $S$  es compacto, tanto  $\varepsilon$  como  $\delta$  son positivos.

Ahora, volvamos al camino dinámico original. Como se muestra arriba, el  $\min_i y_i(t) = m(t)$  y el  $\max_i y_i(t) = M(t)$ , respectivamente, son suficientemente cercanas a  $m^*$  y  $M^*$  para cualquier  $t \geq T$ , provisto  $T$  es tomado suficientemente grande. Entonces, cualquier punto en el camino dinámico es suficientemente cercano al punto en  $S$ . De la continuidad de los  $H^i$ 's.

$$m(t + \tau) - m(t) > \frac{\varepsilon}{2} > 0$$

y

$$M(t) - M(t + \tau) > \frac{\delta}{2} > 0$$

para  $t \geq T$ , provisto  $T$  es suficientemente grande. Pero esto contradice

$$\lim_{t \rightarrow \infty} m(t) = m^* \text{ y } \lim_{t \rightarrow \infty} M(t) = M^*.$$

Por lo tanto,

$$m^* = M^*.$$

Este es el resultado deseado. Esto es notado aquí que todos los componentes del punto inicial  $X(0)$  son no negativos y por lo menos uno de ellos es positivo, entonces esta propiedad se mantiene para cualquier punto  $X(t)$  para todo  $t \geq 0$ .

También es notado aquí que el razonamiento desarrollado arriba no es válido para el sistema de ecuaciones en diferencias

$$X_i(t + 1) = H^i(X_1(t), \dots, X_n(t)), \quad (i = 1, \dots, n).$$

Esto es, si  $H^i$  es creciente en  $X_i$ , podemos construir un ejemplo en el cual el sistema de ecuación en diferencia es inestable. Morishima tiene mostrado la estabilidad relativa del sistema de ecuación en diferencia bajo el supuesto que  $H^i$ 's son no decrecientes en todos los  $X_j$ 's y  $(H^1, \dots, H^n)$  es indescomponible y primitivo, es decir, el supuesto de decrecentabilidad del  $H^i$  en  $X^i$  y la primitividad son adicionalmente requeridas.

La estabilidad es mostrada como nuestro incluso sin la suposición de la primitividad. La indescomposibilidad es suficiente. Pero, aquí nuevamente la estabilidad no es obtenida para el sistema de ecuación en diferencia sin el supuesto de primitividad, esto es, podemos contruir un ejemplo en el cual la inestabilidad es mostrada con la indescomposibilidad pero sin la primitividad. Resumiendo los resultados, la estabilidad es mostrada para el sistema de ecuación diferencial sin los supuestos de no decrecibilidad del  $H^i$  en  $X^i$  y la primitividad.

La razón por qué podemos relajar estos supuestos para el sistema de ecuación diferencial, pero no para el sistema de ecuación en diferencias será explicado en la siguiente sección.  $\square$

## 6. Retorno de escala constante – Caso matricial

Nuestro sistema en el caso es

$$\dot{X} = AX. \quad (46)$$

Aquí,  $X$  es un vector cuyas componentes son los  $X_i$ 's.  $A$  es una matriz indescomponible del cual los elementos de su diagonal son asumidos todos no negativos. Esto es,  $A$  es una matriz Metzler. El siguiente teorma es provisto en esta sección.

**Teorema 2.** Para el sistema de ecuación diferencial bajo la suposición que todos los elementos de su diagonal de  $A$  son no negativos, y  $A$  es indescomponible, existe un único camino del crecimiento balanceado o decaimiento, y cualquier camino solución se aproxima relativamente a este.

Note que el tasa de “crecimiento” puede ser negativo.

**Demostración.** Sea  $\alpha$  un número positivo que es mayor que el valor absoluto de cualquier elemento de la diagonal de la matriz  $A$ . Ponga

$$B \equiv A + \alpha I.$$

Aquí,  $I$  es la matriz identidad. Entonces, todos los elementos de  $B$  son negativos y  $B$  es indescomponible. Entonces,  $B$  tiene un único autovalor positivo  $\mu_1$  y un único autovector positivo  $\bar{X}^{(1)}$  asociado con este tal que  $\mu_1$  no es mayor que los valores absolutos de otros autovalores  $\mu_i$ 's ( $i = 2, \dots, n$ ) de la matriz  $B$ . Ahora, es fácilmente ver que el  $\mu_i - \alpha (\equiv \lambda_i)$  son autovalores de  $A$ . Para

$$\mu_i \bar{X}^{(i)} = B \bar{X}^{(i)} = (A + \alpha I) \bar{X}^{(i)},$$

y además

$$\lambda_i \bar{X}^{(i)} = (\mu_i - \alpha) \bar{X}^{(i)} = A \bar{X}^{(i)}.$$

Aquí,  $\bar{X}^{(i)}$  es el autovector asociado con  $\mu_i$  y  $\bar{X}^{(i)} \not\propto 0$  para  $i \neq 1$ . De arriba, notams que  $\bar{X}^{(i)}$  es un autovector asociado con  $\lambda_i$ , y que  $A$  tiene un único autovector positivo normalizado  $\bar{X}^{(1)}$ . La solución de es escrito explícitamente en la siguiente manera:

$$X(t) = \sum_{i=1}^n c_i \bar{X}^{(i)} e^{\lambda_i t}. \quad (47)$$

(Aquí, este es asumido que cualquier autovalor de una matriz  $A$  tiene un única raíz de la ecuación característica

$$|A - \lambda I| = 0,$$

pero esta suposición no es esencial para la siguiente discusión). Ahora considere los autovalores de  $A + \alpha I$ . El valor absoluto de  $\mu_i$  atrae un máximo cuando  $i = 1$ . Volviendo a llamar  $\mu_1$  es simple, real y positivo, vemos que la parte real de  $\mu_i$  también atrae un máximo cuando y solo cuando  $i = 1$ . Dado

$$\lambda_i = \mu_i - \alpha, \quad (i = 1, \dots, n)$$

vemos que la parte real de  $\lambda_i$  también atrae un máximo cuando y solo cuando  $i = 1$ . Entonces, denotamos de la expresión que la solución de es dominada por el primer término  $c_1 \bar{X}^{(1)} e^{\lambda_1 t}$  en la sumatoria cuando  $t \rightarrow \infty$ . Dado que  $\bar{X}^{(1)}$  es estrictamente positiva, la estabilidad relativa del camino del crecimiento balanceado  $c_1 \bar{X}^{(1)} e^{\lambda_1 t}$  es probado.

Sin embargo tenemos que mostrar que los valores de los  $X_i(t)$ 's se mantienen no negativos provisto las condiciones iniciales de los  $X_i(t)$ 's escogidos así. Esto es fácilmente visto como sigue. Suponga que  $X_1(t) = 0$ . Entonces

$$\begin{aligned} \dot{X}_1(t) &= a_{11}X_1(t) + a_{12}X_2(t) + \dots + a_{1n}X_n(t) \\ &= a_{12}X_2(t) + \dots + a_{1n}X_n(t) \geq 0. \end{aligned}$$

Por lo tanto, la solución del sistema no va en una región con un significado económico donde algunas componentes de  $X$  son negativas. El teorema está probado.  $\square$

Este es almenos el mismo procedimiento como se usó para mostrar el ítem es absolutamente (no relativamente) estable si y solo si el autovalor de la matriz de Metzler con la mayor parte real es negativa. En este sentido, nuestro teorema es solo una extensión trivial de esta propiedad. Citamos el teorema, sin embargo, para explicar el por qué del modelo empleado para probar este teorema no es aplicable al sistema de ecuación en diferencia. Esto es, el sistema

$$X(t+1) = AX(t)$$

no es necesariamente relativamente estable si  $A$  es una matriz de Metzler.

Los valores absolutos de los autovalores son relevantes para la estabilidad del caso ecuación diferencial. En el procedimiento hemos seguido los autovalores de  $A + \alpha I$  para Tan pronto como la parte real es conocida, la posición relativa de los autovalores son mantenidos iguales. Pero, por supuesto el valor absoluto hace cambios. Esto explica por qué la relajación del supuesto de no negatividad de los elementos de la diagonal de  $A$  es posible para el sistema de ecuación diferencial, pero no para el sistema de ecuación en diferencia. El caso no matricial discutivo en la sección precedente también refleja este hecho.

La razón porqué el supuesto de la primitiva es necesario en el caso del sistema de ecuación en diferencia, pero no en el caso del sistema de ecuación diferencial es el mismo. Esto es, los valores absolutos de los autovalores son relevantes para la estabilidad en el caso formado, donde sus partes reales son relevantes para la estabilidad en el último caso.

## 7. Retornos de escala decrecientes

En esta sección, estudiamos el siguiente sistema

$$\dot{X}_i = H^i(X_1, \dots, X_n) \equiv F^i(X_1, \dots, X_n) - \delta_i X_i, \quad (i = 1, \dots, n). \quad (48)$$

Aquí,  $F^i$  es, por ejemplo, la salida del bien capital del tipo  $i$ , y  $\delta_i$  es la tasa de depreciación instantánea del bien capital del tipo  $i$ . Asumamos que todos los  $F^i$ 0s son estrictamente positivo para cualquier  $X$  estrictamente positivo, diferenciable con respecto a cualquier variable y positivamente homogénea de grado  $m$ , los cual son menores que uno, y que

$$\frac{\partial H^i}{\partial X_j} \equiv \frac{\partial F^i}{\partial X_j} \geq 0 \quad \text{para } j \neq i. \quad (49)$$

Aquí, no necesitamos asumir que

$$\frac{\partial H^i}{\partial X_i} \equiv \frac{\partial F^i}{\partial X_i} - \delta_i > 0, \quad (i = 1, \dots, n).$$

y la indescomposibilidad de la matriz  $H_j^i$ . Dado que el propósito principal es mostrar la estabilidad del sistema, asumiremos del conjunto de salida la existencia de la única y equilibrio estrictamente positivo  $(\bar{X}_1, \dots, \bar{X}_n)$ . Esto es notado aquí que incluso si fuéramos a suponer que  $\partial H^i / \partial X_i > 0$  para nuestro sistema, nuestro sistema podría no ser un caso especial de Muth y Suit. Asumimos la homogeneidad de  $F^i$ , pero no  $H^i$ . Es más, incluso si  $\delta_i = 0$  para todo  $i$ , nuestro sistema podría no ser un caso especial de ellos. Para tener asumido que el grado de homogeneidad en un sector puede ser diferente de aquellos en otros sectores. En el caso de Muth, ellos son todos iguales. En el caso de Suit, una forma más general de homogeneidad es introducida, pero el grado de homogeneidad es el mismo en cada sector de producción. Ahora probaremos el siguiente teorema.

**Teorema 3.** Bajo los supuestos de , el grado menor que uno de homogeneidad para todos los  $F^i$ 's y la existencia y unicidad y equilibrio positivo, la solución del sistema ecuación diferencial se aproxima al equilibrio.

**Demostración.** De

$$\frac{\dot{X}_i}{X_i} = \frac{1}{X_i} F^i(X_1, \dots, X_n) - \delta_i, \quad (i = 1, \dots, n). \quad (50)$$

Ponga

$$\frac{1}{X_i} F^i(X_1, \dots, X_n) \equiv G^i(X_1, \dots, X_n), \quad (i = 1, \dots, n). \quad (51)$$

Entonces,  $G^i$  es homogénea de grado  $m_i - 1$  cuyo grado es negativo. Ponga

$$\log X_i = \xi_i, \quad (i = 1, \dots, n). \quad (52)$$

Entonces,

$$X_i = e^{\xi_i} \text{ y } \dot{X}_i/X_i = \dot{\xi}_i, \quad (i = 1, \dots, n).$$

De

$$\dot{\xi}_i = G^i(e^{\xi_1}, \dots, e^{\xi_n}) - \delta_i, \quad (i = 1, \dots, n).$$

Ponga

$$G^i(e^{\xi_1}, \dots, e^{\xi_n}) - \delta_i \equiv g^i(\xi_1, \dots, \xi_n), \quad (i = 1, \dots, n). \quad (53)$$

Entonces,

$$\dot{\xi}_i = g^i(\xi_1, \dots, \xi_n), \quad (i = 1, \dots, n). \quad (54)$$

Ahora,  $G^i(X_1, \dots, X_n)$  es homogénea de grado  $m_i - 1$ . Así,

$$(m_i - 1) G^i = \sum_{j=1}^n \frac{\partial G^i}{\partial X_j} X_j, \quad (i = 1, \dots, n).$$

Dado que  $m_i - 1 < 0$  para todo  $i$ , obtenemos

$$\sum_{j=1}^n \frac{\partial G^i}{\partial X_j} X_j < 0 \quad \text{para todo } i. \quad (55)$$

Ahora calculamos  $\partial g^i / \partial \xi_j$ . De

$$\frac{\partial g^i}{\partial \xi_j} = \frac{\partial G^i}{\partial X_j} \frac{\partial X_j}{\partial \xi_j} = \frac{\partial G^i}{\partial X_j} X_j. \quad (56)$$

Asumimos que

$$\frac{\partial F^i}{\partial X_j} \geq 0 \quad \text{para } j \neq i.$$

Entonces, de

$$\frac{\partial G^i}{\partial X_j} = \frac{\partial}{\partial X_j} \left( \frac{1}{X_i} F^i \right) = \frac{1}{X_i} \frac{\partial F^i}{\partial X_j} \geq 0 \quad \text{para } j \neq i. \quad (57)$$

Por lo tanto, de

$$\frac{\partial g^i}{\partial \xi_j} \geq 0 \quad \text{para } j \neq i. \quad (58)$$

De

$$\frac{\partial G^i}{\partial X_i} X_i < - \sum_{j \neq i} \frac{\partial G^i}{\partial X_j} X_j \leq 0. \quad (59)$$

Entonces, de

$$\frac{\partial g^i}{\partial \xi_i} < 0. \quad (60)$$

De , tenemos

$$\left| \frac{\partial g^i}{\partial \xi_i} \right| > \sum_{j \neq i} \left| \frac{\partial g^i}{\partial \xi_j} \right| \quad \text{para todo } i. \quad (61)$$

Las relaciones son suficientes para la estabilidad del sistema y en consecuencia, el sistema Las relaciones son conocidas como la condición de la diagonal dominantes, y la estabilidad del sistema satisfaciendo esto es mostrado por Arrow, Block and Hurwicz. En la parte superior, asumimos la homogeneidad de las funciones  $F^i(X_1, \dots, X_n)$ ,  $(i = 1, \dots, n)$ . Pero tal suposición no es necesariamente para la estabilidad. Si podemos obtener la relación la estabilidad es obtenida también. Considere el siguiente conjunto de alternativas. Asuma que las cantidades de recursos naturales (incluso la fuerza laboral) son dadas. Sean ellos  $Z_1, \dots, Z_m$ . Asuma que las funciones de producciones

$$F^i(X_1, \dots, X_n, Z_1, \dots, Z_m), \quad (i = 1, \dots, n).$$

Asuma que todos los  $F^i$  son positivamente homogéneas de grado uno en  $X_1, \dots, X_n, Z_1, \dots, Z_m$ . Cuando tomamos en la cuenta todos los tipos de factores de producción, el supuesto del primer grado de homogeneidad es natural. Ahora  $G^i$  es definida en la misma manera como así que  $G^i$  es homogénea de grado cero en  $X_1, \dots, X_n, Z_1, \dots, Z_n$ . Esto es,

$$\sum_{j=1}^n \frac{\partial G^i}{\partial X_j} + \sum_{k=1}^m \frac{\partial G^i}{\partial Z_k} Z_k = 0, \quad (i = 1, \dots, n).$$

Asumiendo que

$$\frac{\partial G^i}{\partial Z_k} \geq 0 \text{ para cada } i \text{ y } k,$$

y que

$$\frac{\partial G^i}{\partial Z_k} > 0 \quad \text{para al menos un } k = k_i, (i = 1, \dots, n).$$

obtenemos

$$\sum_{j=1}^n \frac{\partial G^i}{\partial X_j} X_j < 0, \quad (i = 1, \dots, n).$$

Esto es suficiente para la estabilidad del siguiente sistema,

$$\dot{X}_i = F^i(X_1, \dots, X_n, Z_1, \dots, Z_m) \quad (i = 1, \dots, n).$$

□



## 7.1. Modelo de crecimiento de Solow

Ejemplo 6 (Modelo de crecimiento de Solow). Este modelo de crecimiento neoclásico está basado en la ecuación diferencial

$$\dot{k} = sf(k) - \lambda k \quad (62)$$

Aquí la función desconocida  $k = k(t)$  denota el capital por trabajador,  $s > 0$  denota la tasa constante de ahorro,  $f$  es una función de producción (producto nacional por trabajador como una función del capital por trabajador), y  $\lambda > 0$  denota la tasa proporcional constante de crecimiento del número de trabajadores.

Note que (62) es una ecuación separable. Debido a que  $f$  no se especifica, aún no podemos encontrar una solución explícita de la ecuación. Asuma que el diagrama de fase para la ecuación (62) es como se muestra en la Fig.4. Luego, aquí un estado de equilibrio único con  $k^* > 0$ . Esto es dado por:

$$sf(k^*) = \lambda k^* \quad (63)$$

Por inspección de la Fig.4 vemos que  $k^*$  es estable. Sin importar cuál ha sido el capital inicial por trabajador  $k(0)$ ,  $k(t) \rightarrow k^*$  cuando  $t \rightarrow \infty$ .

Este es un modelo más detallado que lleva a la ecuación (62). Sea  $X(t)$  que denota el ingreso nacional,  $K(t)$  el capital, y  $L(t)$  el número de trabajadores en un país en un tiempo  $t$ . Asuma que

$$\bullet \quad X(t) = F(K(t), L(t)) \quad \bullet \quad \dot{K}(t) = sX(t) \quad \bullet \quad L(t) = L_0 e^{\lambda t}$$

donde  $F$  es una función de producción, y  $s$  es la tasa de ahorro. Asuma que  $F$  es homogénea de grado 1, así que  $F(K, L) = LF(K/L, 1)$  para todo  $K$  y  $L$ . Defina  $k(t) = K(t)/L(t)$  = capital por trabajador, y  $f(k) = F(k, 1) = F(K/L, 1) = F(K, L)/L$  = salida por trabajador. Luego,  $\dot{k}/k = (d/dt)(\ln k) = (d/dt)(\ln K - \ln L)$ , y así

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L} = \frac{sF(K, L)}{K} - \lambda = \frac{sLf(k)}{K} - \lambda = \frac{sf(k)}{k} - \lambda \quad (64)$$

de la cual (62) sigue a la vez.

Observación 1. Déjenos discutir brevemente las condiciones suficientes para la existencia y unicidad del equilibrio del modelo de Solow. Es usual asumir que  $f(0) = 0$ , así como que  $f'(k) > 0$  y  $f''(k) < 0$  para todo  $k > 0$ . Esto es también común postular las llamadas condiciones de Inada, de acuerdo con  $f'(k) \rightarrow \infty$  y también  $f'(k) \rightarrow 0$  cuando  $k \rightarrow \infty$ .

Para ver por qué estas condiciones son suficientes, defina  $G(k) = sf(k) - \lambda k$ . Entonces,  $G'(k) = sf'(k) - \lambda$ , y la ecuación (62) cambia a  $\dot{k} = G(k)$ . Los supuestos sobre  $f$  implica que  $G(0) = 0$ ,  $G'(k) \rightarrow \infty$  cuando  $k \rightarrow 0$ ,  $G'(k) \rightarrow -\lambda < 0$  cuando  $k \rightarrow \infty$ , y  $G''(k) = sf''(k) < 0$  para todo  $k > 0$ . Así  $G$  tiene un único punto estacionario  $\hat{k} > 0$  en el cual  $G'(\hat{k}) = 0$ . Obviamente,  $G(\hat{k}) > 0$ . Pero,  $G'(k) < -\frac{1}{2}\lambda < 0$  para cualquier  $k$  suficientemente grande. Se sigue que  $G(k) \rightarrow -\infty$  cuando  $k \rightarrow \infty$ , así que existe un único punto  $k^* > 0$  con  $G(k^*) = 0$ . Adicionalmente,  $G'(k^*) < 0$ . De acuerdo con esta es una condición suficiente para la estabilidad local asintótica de  $k^*$ .

Las constantes  $\alpha$  y  $\beta$  tiene un significado económico de acuerdo a su valor.

- $\alpha + \beta = 1$ : la función de producción tiene vueltas a escala constante (cambios en la salida subsecuente a un cambio proporcional en las entradas)
- $\alpha + \beta < 1$ : la función de producción tiene vueltas a escala que disminuyen.
- $\alpha + \beta > 1$ : la función de producción tiene vueltas a escala que aumentan.

## 7.2. Deducción algebraica de la función de producción de Cobb-Douglas

Dentro de los supuestos básicos de la función de producción Cobb-Douglas, se tiene:

- Si la mano de obra o capital se reduce, la producción también se reduce en la misma proporción.
- La productividad marginal de la mano de obra es proporcional a la cantidad de producción por unidad de mano de obra.
- La productividad marginal del capital es proporcional a la cantidad de producción por unidad de capital.

Con base a dichas suposiciones, se plantean las ecuaciones diferenciales relacionadas con este comportamiento:

$$\frac{\partial P}{\partial L} = \alpha \frac{P}{L} \quad (65)$$

$$\frac{\partial P}{\partial K} = \beta \frac{P}{K} \quad (66)$$

En relación a las ecuaciones (65) y (66) se puede decir que:

$$K \frac{\partial P}{\partial K} = \beta P \quad (67)$$

$$L \frac{\partial P}{\partial L} = \alpha P \quad (68)$$

Sumando las ecuaciones (67) y (68), sería

$$L \frac{\partial P}{\partial L} + K \frac{\partial P}{\partial K} = \alpha P + \beta P \quad (69)$$

$$L \frac{\partial P}{\partial L} + K \frac{\partial P}{\partial K} = (\alpha + \beta) P \quad (70)$$

Haciendo  $r = \alpha + \beta$ , entonces

$$L \frac{\partial P}{\partial L} + K \frac{\partial P}{\partial K} = r P \quad (71)$$

La ecuación (71) es equivalente al teorema de Euler para funciones homogéneas, lo que indica que si  $r = 1$ , entonces se tendrá una ecuación homogénea de grado 1 y

$$L \frac{\partial P}{\partial L} + K \frac{\partial P}{\partial K} = P(L, K) \quad (72)$$

La ecuación (72) proporciona la productividad marginal de la mano de obra. Como esta ecuación es una ecuación diferencial ordinaria, la solución la hallamos separando variables e integrando. Así, obtenemos

$$\ln(P) + c_1 = \alpha \ln(L) + \beta \ln(K) + c_2. \quad (73)$$

O equivalentemente,

$$\ln(P) = \alpha \ln(L) + \beta \ln(K) + C \quad (74)$$

$$P = e^{\ln(L)^\alpha} e^{\ln(K)^\beta} e^C \quad (75)$$

Haciendo  $A = e^C$  y  $h(K) = e^{g(K)}$  la ecuación (75) se transforma:

$$P = AL^\alpha h(K). \quad (76)$$

Se sabe que:

$$\frac{\partial P}{\partial K} = \beta \frac{P}{K} \quad (77)$$

Derivando parcialmente la función encontrada en el procedimiento anterior y reemplazando:

$$\frac{\partial P}{\partial K} = AL^\alpha h(K) \quad (78)$$

$$\beta \frac{P}{K} = AL^\alpha h(K) \quad (79)$$

$$\beta \frac{AL^\alpha h(K)}{K} = AL^\alpha h(K) \quad (80)$$

La cual se convierte en una ecuación diferencial ordinaria:

$$h'(K) - \beta \frac{h(K)}{K} = 0. \quad (81)$$

La solución de esta ecuación diferencial es  $h(K) = K^\beta$ . Lo cual se verifica fácilmente, ya que al reemplazar en la ecuación anterior se obtiene una identidad. Luego,

$$h(K) = K^\beta \quad (82)$$

$$h'(K) = \beta K^{\beta-1} \quad (83)$$

Reemplazando en la ecuación (81)

$$\begin{aligned} \beta K^{\beta-1} - \frac{\beta K^\beta}{K} &= 0 \\ \frac{\beta K^\beta}{K} &= \frac{\beta K^\beta}{K} \end{aligned}$$

Realizando la sustitución  $y = h(K)$  se tiene  $\frac{dy}{dK} = h'(K)$ .

Reemplazando en la ecuación (81)

$$\frac{dy}{dK} - \beta \frac{y}{K} = 0 \quad (84)$$

Separando variables e integrando obtenemos,

$$\ln(y) + c_1 = \beta \ln(K) + c_2 \quad (85)$$

$$\begin{aligned} \ln(y) &= \beta \ln(K) + c \iff e^{\ln(y)} = e^{(\beta \ln(K) + c)} \iff y = e^{(\ln(K)^\beta + c)} \\ y &= K^\beta e^c \iff y = cK^\beta \\ h(K) &= cK^\beta \end{aligned}$$

Luego de encontrar  $h(K)$ , tenemos que

$$b = Ac, \quad P = bL^\alpha h(K) \iff P = bL^\alpha K^\beta$$

Y dentro de la suposición de la función de producción se sabe que  $\alpha + \beta = 1$ , por lo tanto

$$P = bL^\alpha K^{1-\alpha} \quad (86)$$

### 7.3. Regresión lineal de la función de Cobb-Douglas para un caso en particular

Para poder hallar los valores de  $\alpha$ ,  $\beta$  y  $b$  se debe linealizar la función de producción.

La ecuación (86) la podemos escribir como  $\frac{P}{K} = bL^\alpha K^{-\alpha}$  de donde

$$\begin{aligned} \ln\left(\frac{P}{K}\right) &= \ln\left(b\left(\frac{L}{K}\right)^\alpha\right) \iff \ln\left(\frac{P}{K}\right) = \ln(b) + \alpha \ln\left(\frac{L}{K}\right) \\ \ln\left(\frac{P}{K}\right) &= \ln\left(\frac{P}{K}\right) = \ln(b) + \alpha \ln\left(\frac{L}{K}\right) \end{aligned}$$

## 8. El modelo de Solow sobre el crecimiento

- Desarrollar un simple marco para los causes próximos y la mecánica del crecimiento económico y los países cruzados diferencias.
- El modelo de Solow-Swan llamado después de Robert Solow y Trevor Swan.
- Antes del modelo de Solow, la aproximación más común al crecimiento económico se construyó sobre el modelo de Harrod-Domar.
- El modelo de Harrod-Domar enfatizó los potenciales aspectos disfuncionales del crecimiento: por ejemplo, cómo el crecimiento podría ir mano a mano con el incremento del desempleo.
- El modelo de Solow demostró por qué el modelo de Harrod-Domar no fue un lugar atractivo para iniciar.
- En el centro del modelo de crecimiento de Solow está la función de producción agregada neoclásica.
- La economía cerrada, con un único final bueno.
- El tiempo discreto corriendo en un horizonte infinito, el tiempo es indexado por  $t = 0, 1, 2, \dots$
- La economía es inhabitada por un largo número de , y por ahora los no serán optimizados.
- Para fijar ideas, asuma que los son idénticos, así la economía admite un representativo
- Asuma que los hogares guardan una fracción exógena constante  $s$  de tu ingreso disponible.
- Asuma que todas las formas tienen acceso a la misma función de producción: la economía admite una firma representativa, con un representante (o agregado) de la función de producción.
- La función de producción agregada para el único final bueno es

$$Y(t) = F[K(t), L(t), A(t)]. \quad (87)$$

- Asuma que el capital es el mismo como el producto final de la economía, pero usado en el proceso de producción de más productos.
- $A(t)$  es un cambio de la función de producción (87). Noción amplia de la tecnología.
- Suposición principal: la tecnología es gratuita, este está disponible como no excluible, no tiene un producto rival.

Suposición 1 (Continuidad, diferenciabilidad, positiva y productos marginales decrecientes, y retornos de escala constante). La función de producción  $F: \mathbb{R}_+^3 \rightarrow \mathbb{R}$  es dos veces diferenciable en  $K$  y  $L$  y satisface

$$\begin{aligned} F_K(K, L, A) &\equiv \frac{\partial F}{\partial K} > 0, & F_L(K, L, A) &\equiv \frac{\partial F(\cdot)}{\partial L} > 0, \\ F_{KK}(K, L, A) &\equiv \frac{\partial^2 F(\cdot)}{\partial K^2} < 0, & F_{LL}(K, L, A) &\equiv \frac{\partial^2 F(\cdot)}{\partial L^2} < 0. \end{aligned}$$

Es más,  $F$  exhibe un retorno de escala constante en  $K$  y  $L$ .

Asuma que  $F$  exhibe un retorno a escala constante en  $K$  y  $L$ , es decir, es homogénea linealmente (homogénea de grado 1) en estos dos variables.

Definición 4. Sea  $K$  un entero. La función  $g: \mathbb{R}^{K+2} \rightarrow \mathbb{R}$  es homogénea de grado  $m$  en  $x \in \mathbb{R}$  e  $y \in \mathbb{R}$  si y solo si

$$g(\lambda x, \lambda y, z) = \lambda^m g(x, y, z) \text{ para todo } \lambda \in \mathbb{R}_+ \text{ y } z \in \mathbb{R}^K.$$

Teorema 4 (Euler). Suponga que  $g: \mathbb{R}^{K+2} \rightarrow \mathbb{R}$  es continuamente diferenciable en  $x \in \mathbb{R}$  e  $y \in \mathbb{R}$ , con derivadas parciales denotada por  $g_x$  y  $g_y$  y es homogénea de grado  $m$  en  $x$  e  $y$ . Entonces

$$mg(x, y, z) = g_x(x, y, z)x + g_y(x, y, z)y \quad \forall x \in \mathbb{R}, y \in \mathbb{R} \text{ y } z \in \mathbb{R}^K.$$

Más aún,  $g_x(x, y, z)$  y  $g_y(x, y, z)$  son por sí mismo homogéneas de grado  $m - 1$  en  $x$  e  $y$ .

### 8.1. Estructura del mercado, dotaciones y limpieza del mercado

- Asumiremos que los mercados son competitivos, así que nuestro será un prototipo del equilibrio de un modelo general competitivo.
- Los hogares poseen todo el trabajo, que suministran inelásticamente.

```
18
abc
2
x = 55
y = 0
z = 5
```

A message from Python!

```
>>> x_1 = 1 + 1
>>> x_1
2
```

## A. Seleccionar una medida de desempeño

El siguiente paso es seleccionar una medida de desempeño. Una forma típica de medir para problemas de regresión es el error de la raíz media cuadrática (RMSE). Este nos da una idea cómo el error del sistema típicamente hace en sus predicciones, con un alto peso para errores grandes. La ecuación (88)

$$\text{RMSE}(\mathbf{X}, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(\mathbf{x}^{(i)}) - y^{(i)})^2} \quad (88)$$

- $m$  es el número de instancias en el conjunto de datos que se está midiendo.
- $\mathbf{x}^{(i)}$  es un vector de todos los valores de la característica (excluyendo la etiqueta) de la  $i$ -ésima instancia en un conjunto de datos, e  $y^{(i)}$  es su etiqueta (el valor deseado de salida para esa instancia).
- $\mathbf{X}$  es una matriz que contiene todos los valores característicos (excluyendo etiquetas) de todas las instancias en un conjunto de datos.
- $h$  es la función del sistema predictivo, también llamado hipótesis. Cuando el sistema es dado una característica de instancia, su salida es el valor predecido  $\hat{y}^{(i)} = h(\mathbf{x}^{(i)})$  para la instancia.
- $\text{RMSE}(\mathbf{X}, h)$  es la función de costo medido en un conjunto de ejemplos usando la hipótesis  $h$ .

Incluso pensando que la RMSE es generalmente la medida de desempeño preferido para las tareas de regresión, en algunos contextos podría preferir usar otra función. Por ejemplo, suponga que existen muchos distritos outliers. En este caso, podría considerar usar el error cuadrático medio (también llamada la desviación media absoluta, vea la ecuación (89))

$$\text{MAE}(\mathbf{X}, h) = \frac{1}{m} \sum_{i=1}^m |h(\mathbf{x}^{(i)}) - y^{(i)}| \quad (89)$$

Tanto la RMSE como la MAE son maneras de medir la distancia entre dos vectores: el vector de predicción y el vector de valores objetivo. Varias medidas de distancias, son posibles:

- Calculando la raíz cuadrada de una suma de cuadradas (RMSE) corresponde a la norma euclidiana: esta es la noción de distancia con la que está familiarizado. Este es llamado la norma  $\ell_2$ , denotado por  $\|\cdot\|_2$  (o solo  $\|\cdot\|$ ).
- Calculando la suma de los valores absolutos (MAE) corresponde a la norma  $\ell_1$ , denotado por  $\|\cdot\|_1$ . A veces llamada norma Manhattan porque este mide la distancia entre dos puntos en una ciudad si solo puede viajar a lo largo de cuerdas ortogonales.
- Más generalmente, la norma  $\ell_k$  de un vector  $\mathbf{v}$  que contiene  $n$  elementos es definido por  $(|v_0|^k + |v_1|^k + \dots + |v_n|^k)^{\frac{1}{k}}$ .  $\ell_0$  da el número de elementos no nulos en el vector y  $\ell_\infty$  da el máximo valor absoluto en el vector.
- El mayor índice de la norma, se centra en valores grandes y Este es la razón por la que RMSE es más sensitiva a los outliers que el MAE. Pero cuando

En este capítulo, empezaremos mirando un modelo de regresión lineal, uno de los modelos más simple que hay. Discutiremos dos maneras muy diferentes de tratar:

- Usando la fórmula cerrada que directamente calcula los parámetros del modelo que minimiza la función de costo sobre el conjunto de datos.
- Usando un método de optimización iterativa, llamado el descenso del gradiente, que gradualmente ajusta los parámetros para minimizar la función de costo sobre el conjunto de datos, eventualmente convergiendo al mismo conjunto de parámetros como el primer método. Veremos algunas pocas variantes del descenso del gradiente.

Luego, veremos la regresión polinomial, un modelo complejo que puede ajustar conjunto de datos no lineales. Dado que este modelo tiene más parámetros que la regresión lineal, este es así veremos cómo detectar cuando es o no el caso, usando curvas de aprendizaje, y entonces veremos varias técnicas de regularización que pueden reducir el sobreajuste del conjunto de datos. Finalmente, veremos sobre dos modelos comúnmente usados para tareas de clasificación: la regresión logística y la regresión softmax.

En Este modelos es solo una función lineal con características de entrada  $\theta_0$  y  $\theta_1$  son los parámetros del modelo.

Más generalmente, un modelo lineal hace una predicción por simple cálculo de suma de pesos de características de entrada, más una constante llamada el término intercepto, como se muestra en la ecuación (90)

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots \theta_n x_n \quad (90)$$

- $\hat{y}$  es el valor predicho.
- $n$  es el número de características.
- $\theta_j$  es el  $j$ -ésimo parámetro del modelo (incluyendo el término intercepto  $\theta_0$  y los pesos de las características  $\theta_1, \theta_2, \dots, \theta_n$ ).

Esto puede ser escrito mucho más conciso usando una forma vectorial, como se muestra en (91)

$$\hat{y} = h_{\theta}(x) = \theta \cdot x \quad (91)$$

- $\theta$  es parámetro vector del modelo, conteniendo el término intercepto  $\theta_0$  y los pesos características desde  $\theta_1$  hasta  $\theta_n$ .
- $x$  es la instancia del vector característica, conteniendo desde  $x_0$  hasta  $x_n$ , con  $x_0 = 1$ .
- $\theta \cdot x$  es el producto interno de  $\theta$  y  $x$ , el cual es igual a  $\theta_0 x_0 + \theta_1 x_1 + \cdots \theta_n x_n$ .
- $h_{\theta}$  es la función de hipótesis, usando los parámetros  $\theta$  del modelo.

En el apéndice 1 vimos que la forma más común de medir el desempeño de un modelo de regresión es la raíz cuadrática media (RMSE). Por lo tanto, para emplear el modelo de regresión lineal, necesitarás encontrar el valor de  $\theta$  que minimice la RMSE. En la práctica, es más simple minimizar el error cuadrático medio (MSE) que el RMSE, y se consigue el mismo resultado (porque el valor que minimiza una función también minimiza su raíz cuadrada).

El MSE de una hipótesis de regresión lineal  $h_{\theta}$  en un conjunto de datos  $\mathbf{X}$  es calculado usando la ecuación (92)

$$\text{MSE}(\mathbf{X}, h_{\theta}) = \frac{1}{m} \sum_{i=1}^m \left( \theta^T x^{(i)} - y^{(i)} \right)^2 \quad (92)$$

La única diferencia es que escribimos  $h_{\theta}$  en vez de solo  $h$  para hacer más claro que el modelo es parametrizado por el vector  $\theta$ . Para simplificar notaciones, solo escribiremos  $\text{MSE}(\theta)$  en vez de  $\text{MSE}(\mathbf{X}, h_{\theta})$ .

## A.1. La ecuación normal

Para encontrar el valor de  $\theta$  que minimice la función de costo, existe una solución en forma cerrada, en otras palabras, una ecuación matemática que nos da el resultado directo. Esto es llamado la ecuación normal

$$\hat{\theta} = (X^T X)^{-1} X^T y \quad (93)$$

- $\hat{\theta}$  es el valor de  $\theta$  que minimiza la función de costo.
- $y$  es el vector de valores objetivos conteniendo desde  $y^{(1)}$  hasta  $y^{(m)}$ .

Ahora generemos datos para probar esta ecuación en

```
>>> import numpy as np
>>> X = 2*np.random.rand(100, 1)
>>> y = 4 + 3*X + np.random.randn(100, 1)
```

Ahora calculemos  $\hat{\theta}$  usando la ecuación normal. Usaremos la función `inv()` del módulo de álgebra lineal de Numpy (`np.linalg`) para calcular la inversa de una matriz, y el método `dot()` para la multiplicación de matrices:

```
>>> X_b = np.c_[np.ones((100, 1)), X] # Sumar x0 = 1 para cada instancia
>>> theta_best = np.linalg.inv(X_b.T.dot(X_b)).dot(X_b.T).dot(y)
```

La función actual usaremos para generar este dato es  $y = 4 + 3x_1 + \text{Ruido gaussiano}$ . Vemos que la ecuación encontrada:

```
>>> theta_best
array([[4.22606177],
       [2.92965516]])
```

Podríamos esperar para  $\theta_0 = 4$  y  $\theta_1 = 3$  en vez de  $\theta_0 = 4.215$  y  $\theta_1 = 2.770$ . Muy cercano, pero el ruido hace imposible recuperar los parámetros exactos de la función original.

Ahora puede hacer predicciones usando  $\hat{\theta}$ :

```
>>> X_new = np.array([[0], [2]])
>>> X_new_b = np.c_[np.ones((2, 1)), X_new] # Suma x0=1 en cada instancia
>>> y_predict = X_new_b.dot(theta_best)
>>> y_predict
array([[ 3.86893532],
       [10.18025405]])
```

Ahora grafiquemos los modelos de predicciones (): Mejoramos la regresión lineal usando Scikit-Learn es un poco simple:

```
>>> from sklearn.linear_model import LinearRegression
>>> lin_reg = LinearRegression()
>>> lin_reg.fit(X, y)
>>> lin_reg.intercept_, lin_reg.coef_
(array([4.21509616]), array([2.77011339]))
>>> lin_reg.intercept_, lin_reg.coef_
array([4.21509616], [9.75532293])
```

La clase `LinearRegression` está basado en la función `scipy.linalg.lstsq()` (el nombre abreviado de “mínimos cuadrados”), el cual puede llamarlo directamente:

```
>>> theta_best_svd, residuals, rank, s = np.linalg.lstsq(X_b, y, rcond=1e-6)
>>> theta_best_svd
array([4.21509616], [2.77011339])
```

La función calcula  $\hat{\theta} = \mathbf{X}^+ \mathbf{y}$ , donde  $\mathbf{X}^+$  es la pseudoinversa de  $\mathbf{X}$  (específicamente la inversa de Moore-Penrose). Puede usar `np.linalg.pinv()` para calcular la pseudoinversa directamente:

```
>>> np.linalg.pinv(X_b).dot(y)
array([4.21509])
```

La pseudoinversa por sí misma es calculada usando la técnica estándar de factorización de matrices llamada la descomposición de valores singulares que puede ser descompuesta la matriz  $\mathbf{X}$  en una multiplicación de tres matrices  $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  (vea `numpy.linalg.svd()`). La pseudoinversa es calculada como  $\mathbf{X}^+ = \mathbf{V}\mathbf{\Sigma}^+ \mathbf{U}^T$ . Para calcular la matriz  $\mathbf{\Sigma}^+$ , el algoritmo toma  $\mathbf{\Sigma}$  y fija a cero todos los valores menores que un pequeño, entonces se reemplaza todos los valores distintos de cero con su inversa, y finalmente se transpone la matriz resultante. Esta aproximación es más eficiente que calcular la ecuación normal, más es más, la ecuación normal podría no trabajar si la matriz  $\mathbf{X}^T \mathbf{X}$  no es inversible (es decir, singular), así como si  $m < n$  o si alguna de sus características son redundantes, pero la pseudoinversa está siempre definida.

## A.2. Complejidad computacional

La ecuación normal calcula la inversa de  $\mathbf{X}^T \mathbf{X}$ , que es una matriz  $(n+1) \times (n+1)$  (donde  $n$  es el número de características), La complejidad computacional de la inversión de tal matriz es típicamente acerca de  $\mathcal{O}(n^{2.4})$  hasta  $\mathcal{O}(n^3)$  (dependiendo en la implementación). En otras palabras, si dobla el número de características, multiplique el tiempo de cálculo por  $2^{2.4} = 5.3$  hasta  $2^3 = 8$ .

El enfoque SVD usado por la clase `LinearRegression` por Scikit-Learn es acerca  $\mathcal{O}(n^2)$ . Si dobla el número de características, multiplica el tiempo de cálculo hasta por 4.

También, una vez que los datos estén en el modelo de regresión lineal (usando la ecuación normal o cualquier otro algoritmo), las predicciones son muy rápidas: la complejidad comutacional es lineal con

Ahora, vemos otras maneras diferentes de emplear el modelo de regresión lineal,



### A.3. Descenso del gradiente

El descenso del gradiente es un algoritmo de optimización muy genérico para encontrar soluciones óptimas a un amplio rango de problemas. La idea general del descenso del gradiente es para mejorar los parámetros iterativamente con el fin de minimizar la función de costo.

Suponga que está perdido en las montañas en una densa niebla, puede solo sentir la pendiente del suelo bajo sus pies. Una buena estrategia es conseguir el fondo del valle rápidamente hacia en la dirección de pendiente del suelo. Este es exactamente lo que el descenso del gradiente hace: este mide el gradiente local de la función error con del parámetro vectorial  $\theta$ , y va en la dirección del descenso del gradiente. Una vez que el descenso del gradiente es cero, ¡ya has alcanzado un mínimo!

Concretamente, empieza por completar  $\theta$  con valores aleatorias (este es llamado iniciación aleatoria), y entonces mejoras gradualmente, tomando un pequeño paso por vez, cada paso intenta decrecer la función de costo (por ejemplo, el MSE), bajo la convergencia del algoritmo a un mínimo.

Un parámetro importante en el descenso del gradiente es el tamaño de los pasos, determinado por el hiperparámetro tasa de aprendizaje. Si la tasa de aprendizaje es muy pequeña, entonces el algoritmo tiene que pasar muchas iteraciones para converger, el cuál podría tomar un largo tiempo.

Por otro lado, si la tasa de aprendizaje es muy alta, podría saltar a lo largo del valle hasta el fin del lado opuesto, posiblemente más alto de donde estuvo antes. Esto podría hacer que el algoritmo diverga, con valores más grandes, fallando en la búsqueda de una buena solución.

Finalmente, no todas las funciones costo lucen como una suave Podría haber agujeros, riscos y todo tipo de terrenos irregulares, haciendo la convergencia al mínimo muy difícil. Muestra los dos retos principales con el descenso del gradiente: si la inicialización aleatoria empieza con el algoritmo en la izquierda, entonces convergerá a un mínimo local, que no es tan bueno como el mínimo global. Si este empieza por la derecha, entonces este tomará un largo tiempo a la platea, y si te detienes muy pronto no alcanzarás el mínimo global.

Fortunamente,

Facultad de Ciencias, 21 de octubre del 2019.

---

\*Email: [kfernandezh@uni.pe](mailto:kfernandezh@uni.pe)

†Email: [aaberrospic@uni.pe](mailto:aaberrospic@uni.pe)

‡Email: [btorresa@uni.pe](mailto:btorresa@uni.pe)

§Email: [caznaranl@uni.pe](mailto:caznaranl@uni.pe)

¶[ctorresp@uni.edu.pe](mailto:ctorresp@uni.edu.pe)



Universidad Nacional de Ingeniería  
Facultad de Ciencias  
Av. Túpac Amaru N° 210 Rimac  
Lima, Perú

Presentación teórica del tema

## La función de producción Cobb-Douglas

La función de producción: Cobb-Douglas

K. Fernández Huidobro*	2018228G
A. Berrospi C.†	20182720C
B. Torres Ayalas‡	201721339G
C. Aznarán Laos§	20162720C

October 21, 2019

Supervisado por M.Sc. Clifford Torres Ponce¶

### Abstract

The function of production Cobb-Douglas has a neoclassic focus to estimate the function of production of a country and to project the expected economic growth.

To represent the relationships between the production obtained, it uses the variations of capital ( $K$ ) and labor ( $L$ ) inputs, to which technology was later added, also called total factor productivity ( $PTF$ ). It is a production function frequently used in Economics.

The origin of the Cobb Douglas function is found in the empirical observation of the distribution of the total national income of the United States between capital and labor. According to what the data showed, the distribution remained relatively constant over time. Specifically, work took 70% and capital 30%. In this way, the Cobb Douglas function represents a relationship in which the proportions of labor and capital with respect to the total product are constant.

---

\*Email: [kfernandezh@uni.pe](mailto:kfernandezh@uni.pe)

†Email: [aaberrospic@uni.pe](mailto:aaberrospic@uni.pe)

‡Email: [btorresa@uni.pe](mailto:btorresa@uni.pe)

§Email: [caznaranl@uni.pe](mailto:caznaranl@uni.pe)

¶[ctorresp@uni.edu.pe](mailto:ctorresp@uni.edu.pe)

## Contents