Notas del Curso

Optimización Dinámica Aplicada a Finanzas

Parte Dinámica

Germán Daniel Fermo, Ph.D

D- Optimización Intertemporal

Ecuaciones en Diferencias Lineales de Primer Orden

Una ecuación en diferencias de primer orden exhibe la siguiente forma matemática:

$$y_{t+1} = c - ay_t$$

Es decir, como la ecuación de arriba lo muestra, una ecuación en diferencias es aquélla en donde el valor futuro de la variable (en este caso y) depende del valor pasado de esa misma variable. En el caso particular de ecuaciones en diferencias de primer orden, el valor de la variable depende del valor pasado correspondiente al momento inmediatamente precedente. Es por esta razón, que en una ecuación en diferencias de primer orden, sólo aparecen involucrados los subíndices: $t+1$ y t.

En general, y esto lo vamos a ver muy bien luego, la resolución de todo sistema dinámico implica hallar dos conceptos: 1) el equilibrio de largo plazo de la variable (llamado normalmente por economistas como el steady state), 2) la desviación intertemporal que la variable muestra a lo largo del tiempo respecto de su steady state. Cuando la trayectoria de y es convergente, la desviación intertemporal decrece a lo largo del tiempo, cuando por el contrario la trayectoria de y es divergente, la evolución dinámica de la variable es explosiva y nunca converge al steady state.

Qué quiere decir resolver una ecuación en diferencias?. Resolver una ecuación en diferencias implica encontrar ***TODA*** la trayectoria de la variable y a lo largo del tiempo dependiendo ***EXCLUSIVAMENTE*** del valor inicial de la variable y y de los parámetros de la ecuación. Para resolver la ecuación en diferencias se necesitan encontrar los dos componentes que definimos en el párrafo anterior.

El primer componente se denomina ***integral particular*** y el segundo ***función complementaria***.

Integral Particular

La resolución de la integral particular implica encontrar el steady state del sistema. El sistema alcanza el steady state sí y sólo sí:

$$y_{t+1} = y_t \Rightarrow \Delta y = 0$$

Es decir, cuando la variable ya no varía mas, se dice que se estabilizó en su steady state, es decir, encontró un estado estacionario. Volvamos entonces a la ecuación en diferencias y encontremos el steady state:

$$Si\ y_{t+1} = y_t \Rightarrow \bar{y} = c - a\bar{y} \Rightarrow \bar{y}(1+a) = c \Rightarrow \bar{y} = c/(1+a)$$

Entonces, sabemos que el sistema converge al steady state: $c/(1+a)$.

Función Complementaria

Este componente de la solución de la ecuación en diferencias nos indica el desvío temporal que la variable tiene respecto de su equilibrio de largo plazo (steady state). Resolver esta parte de una ecuación en diferencias es como resolver una integral, hay que hacer un guess de la solución. Probemos con la siguiente expresión:

$$y_t = A(-a)^t \Rightarrow y_{t+1} = A(-a)^{t+1}$$

La solución de la función complementaria requiere encontrar un valor tal que la siguiente condición se cumpla:

$$y_{t+1} = -ay_t$$

Como vemos, si $y_t = A(-a)^t$ tenemos que:

$$y_{t+1} = -ay_t \Rightarrow A(-a)^{t+1} = -aA(-a)^t = A(-a)^{t+1}$$

Vemos que la condición se satisface. Ahora bien, para encontrar finalmente la solución a la ecuación en diferencias necesitamos encontrar el valor del parámetro A. Es por eso que ***TODA*** solución de una ecuación en diferencias de primer orden ***REQUIERE*** de una condición inicial. Es decir, debo conocer el valor de la variable hoy: y_0. Recordemos que la solución de la ecuación en diferencias requiere que se sume la integral particular y la función complementaria. Si conozco ese valor podemos usar la siguiente ecuación para despejar A:

$$y_t = A(-a)^t + c/(1+a) \Rightarrow y_0 = A(-a)^0 + c/(1+a) = A + c/(1+a) \Rightarrow y_0 - c/(1+a) = A$$

Ahora entonces la solución de la ecuación en diferencias es:

$$y_t = A(-a)^t + c/(1+a) \Rightarrow y_t = [y_o - c/(1+a)](-a)^t + c/(1+a)$$

Vemos que la solución de la ecuación en diferencias es encontrar la **TRAYECTORIA TEMPORAL** de la variable expresada ***NO EN FUNCION DE SI MISMA*** sino de los parámetros del modelo (en este caso a y c) y de su valor inicial (y_0). Es ahora en donde se pone interesante el análisis.

Como vemos, toda la dinámica del modelo depende del valor del parámetro "a". En primer lugar, si a es igual a -1, el modelo está indefinido, dado que la división por cero no existe. La dinámica del modelo como función de a puede resumirse en la siguiente tabla:

	Abs(a)<1	**Abs(a)>1**
a>0	Converge, Oscilante	Diverge, Oscilante
a<0	Converge, No Oscilante	Diverge, No Oscilante

Curva de Fase

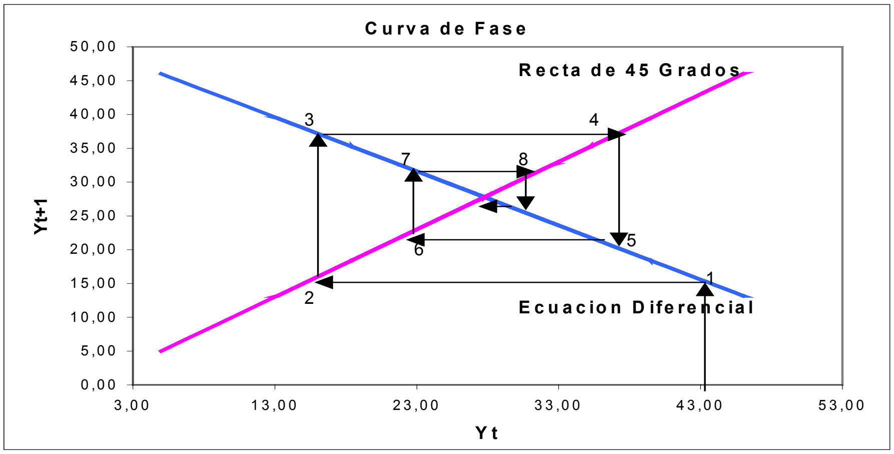
En lugar de encontrar una solución reducida a la ecuación en diferencias, existe un procedimiento más intuitivo (y obviamente menos exacto) denominado curvas de fase. La curva de fase consiste en graficar cómo evoluciona dinámicamente la variable respecto de sí misma. Es decir, en la ordenada se grafica y_{t+1} y en la abscisa se grafica y_t. Para el caso de ecuaciones en diferencias de primer orden **LINEALES** la curva de fase utiliza una recta de 45 grados que parte del origen. Recordemos que si una recta tiene 45 grados y parte del origen equivale a tener una recta con pendiente igual a uno y una ordenada al origen igual a cero. Una recta con esta particularidad funciona como un espejo que me permite mapear la variable y (cualquiera que fuese) desde la ordenada a la abscisa. Entonces, si y_{t+1} en el periodo $t+1$ está en la ordenada, utilizando la recta de 45 grados, en el periodo $t+2$ estará en el eje de abscisas y de esta forma podré utilizar a y_{t+1} como variable independiente en el periodo $t+2$ que me permitirá a su vez calcular y_{t+2} como variable dependiente del periodo $t+2$. A su vez, en el periodo $t+3$ utilizando nuevamente la recta de 45 grados, puedo mapear a y_{t+2} desde la ordenada a la abscisa y utilizarla como variable independiente y determinar de esta forma a y_{t+3}, y así sucesivamente. La relación entre y_{t+1} y y_t (para todo t) se determina utilizando la respectiva ecuación diferencial lineal:

$$y_{t+1} = c - ay_t$$

Volviendo ahora al análisis de estabilidad, analicemos dos casos:

Caso 1:Abs(a)<1 y a>0

Sabemos que en este caso el sistema es ***ESTABLE***. Veamos cómo podemos determinar la estabilidad a través del uso de la curva de fase:

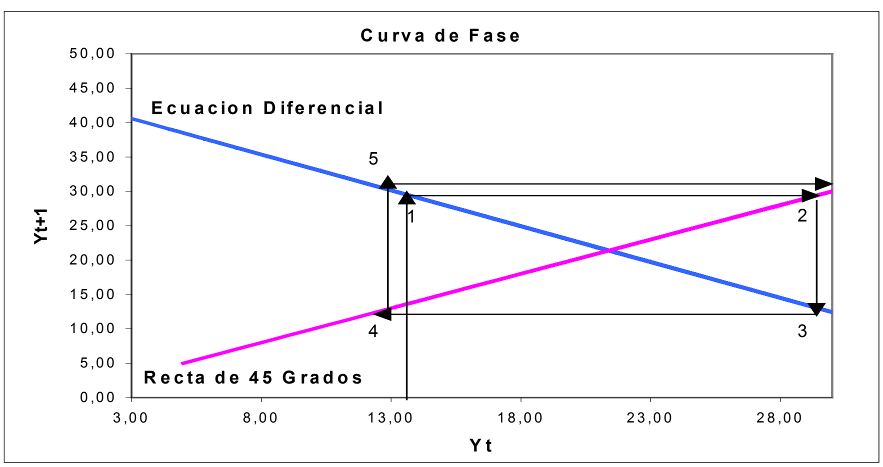


El 1 indica adonde empezó el sistema. Como vemos el diagrama de fase utiliza dos rectas: la de 45 grados que simplemente me permite trasladar a la variable dependiente al eje de las abscisas (lo que la convierte en variable independiente en el periodo siguiente). Es importante destacar que en $t+1$, y_{t+1} es la variable dependiente. Sin embargo, en $t+2$, y_{t+1} se convierte en la variable independiente, y así sucesivamente. Obviamente, una vez que se mapea la variable en la abscisa, se utiliza la ecuación en diferencias para determinar los valores resultantes de y_{t+1}, y_{t+2}, …. etc..

Vemos que en el gráfico precedente, como $abs(a)$ es menor que 1, el sistema converge a un equilibrio estacionario.

Caso 1:Abs(a)>1 y a>0

Sabemos que en este caso el sistema es ***INESTABLE***.



Vemos que aquí el sistema es explosivo y diverge del equilibrio.

Modelo de la Telaraña

Este modelo describe una situación en la que la decisión de producir debe ser hecha en un periodo anterior a la venta del producto. Un ejemplo de esto puede ser la producción agrícola en donde la siembra debe preceder en un desfase significativo de tiempo a la cosecha y posterior venta del producto. Supongamos entonces que la decisión de producción en el periodo t se basa en el precio vigente en t, es decir, P_t. Sin embargo, como esta producción no estará disponible para su venta hasta en periodo siguiente, P_t determinará Q_{t+1} (es decir, la oferta del bien) y no Q_t. Por lo tanto, este modelo se caracteriza por tener una función de oferta ***DESFASADA*** un periodo:

$$Q_t^S = S(P_{t-1})$$

La demanda se comporta como cualquier demanda standard, es decir, dependiendo del precio actual.

$$Q_t^D = D(P_t)$$

Dadas estas dos ecuaciones, para que el sistema quede completo, es necesario introducir una condición de equilibrio:

$$Q_t^S = Q_t^D$$

Ahora bien, supongamos que el modelo exhibe funciones de demanda y oferta lineales determinadas por los siguientes parámetros:

$$Q_t^S = -\lambda + \delta P_{t-1}\ (\lambda, \delta > 0) \quad Q_t^D = \alpha - \beta P_t\ (\alpha, \beta > 0)$$

La condición de equilibrio implica lo siguiente:

$$-\lambda + \delta P_{t-1} = \alpha - \beta P_t \qquad \text{A}$$

Resolviendo esta ecuación llegamos a la siguiente expresión:

$$P_t = (\alpha + \lambda)/\beta - (\delta/\beta)P_{t-1}$$

Vemos que esta ecuación es una ecuación en diferencias de primer orden exactamente igual a la que analizamos precedentemente:

$$y_{t+1} = c - ay_t$$

En donde,

$$c = (\alpha + \lambda)/\beta \quad a = \delta/\beta \Rightarrow (1+a) = (\beta + \delta)/\beta$$

Recordemos que el estado estacionario era:

$$\overline{y} = c/(1+a) \quad \overline{P} = [(\alpha + \lambda)/\beta][\beta/(\beta + \delta)] \Rightarrow \overline{P} = (\alpha + \lambda)/(\beta + \delta)$$

La solución de la ecuación diferencial era:

$$y_t = [y_o - \frac{c}{(1+a)}](-a)^t + \frac{c}{(1+a)} \Rightarrow P_t = [P_0 - \frac{(\alpha + \lambda)}{(\beta + \delta)}](-\frac{\delta}{\beta})^t + \frac{(\alpha + \lambda)}{(\beta + \delta)}$$

Es fundamental resaltar que aun cuando los precios no converjan a su steady state, las cantidades demandadas y ofrecidas en cada periodo son iguales. Esto no es una casualidad. Si observamos la ecuación "A" vemos que la ecuación diferencial desfasada surge de igualar la oferta a la demanda. Vemos entonces que por propia construcción del modelo las cantidades ofrecidas y demandadas son siempre idénticas. Entonces podríamos preguntarnos por qué el sistema se sigue moviendo hasta alcanzar un steady state. La razón es que aun cuando las cantidades demandadas y ofrecidas sean iguales, dado que la oferta reacciona desfasada, los precios que equilibran la oferta difieren de los que equilibran la demanda. El único punto en el que esta dinámica se frena es en el steady state (y esto sólo ocurre si la condición de estabilidad se cumple). Vemos que el comportamiento dinámico del sistema depende de los parámetros δ y β. Estos parámetros son las pendientes de las curvas de oferta y demanda respectivamente. Dados que todos los parámetros son positivos la condición de valor absoluto se reduce a:

	$\delta/\beta<1 \Leftrightarrow \delta<\beta$	$\delta/\beta>1 \Leftrightarrow \delta>\beta$
$\delta/\beta>0 \Leftrightarrow \delta>0$	Converge, Oscilante	Diverge, Oscilante
$\delta/\beta<0 \Leftrightarrow \delta<0$	Converge, No Oscilante	Diverge, No Oscilante

Dado que δ y β son ambos positivos, el segundo renglón de la tabla es irrelevante. Con lo cual podemos deducir que el modelo se va a comportar en forma oscilante. Si la pendiente de la curva de oferta (δ) es inferior a la pendiente de la curva de demanda (β), el sistema converge, de lo contrario es explosivo.

Modelo de Equilibrio de Mercado

Supongamos que, para un determinado bien, las funciones de demanda y oferta son:

$$Q_d = \alpha - \beta P \; (\alpha > 0) \qquad Q_s = -\lambda + \delta P \; (\lambda, \delta > 0)$$

El precio de equilibrio surge de igualar oferta y demanda, resultando en la siguiente expresión:

$$\overline{P} = (\alpha + \lambda)/(\beta + \delta)$$

Para analizar la dinámica, es necesario suponer que existe una ***LEY DE AJUSTE*** de los precios cuando el sistema está en desequilibrio, es decir, cuando la cantidad demandada difiere de la ofrecida. Supongamos entonces que la ley de ajuste es:

$$P_{t+1} - P_t = j(Q_d - Q_s) \; (j > 0)$$

Las implicancias de esta ley de ajuste son muy intuitivas. Dado que $(j > 0)$, cuando la cantidad demandada supera a la ofrecida, los precios suben, de lo contrario, bajan. Es importante analizar el significado de j, este parámetro mide la velocidad con la que los precios se ajustan a desequilibrios en la demanda y oferta.

Si a la ecuación diferencial precedente le introducimos las ecuaciones de demanda y oferta, resulta:

$$P_{t+1} - P_t = j(\alpha - \beta P_t + \lambda - \delta P_t) \Rightarrow P_{t+1} + P_t(-1 + j(\beta + \delta)) = j(\alpha + \lambda)$$

Esto resulta en:

$$P_{t+1} = j(\alpha + \lambda) + (1 - j(\beta + \delta))P_t$$

Como vemos, esta ecuación tiene la misma forma de las ecuaciones en diferencias lineales de primer orden analizadas anteriormente:

$$y_{t+1} = c - a y_t$$

En este caso:

$$a = (-1 + j(\beta + \delta))$$

La solución de la ecuación en diferencias es:

$$P_t = [P_o - (\alpha + \lambda)/(\beta + \delta)](1 - j(\beta + \delta))^t + (\alpha + \lambda)/(\beta + \delta)$$

Recordemos que para que el sistema sea estable se necesita lo siguiente:

$$abs[(1-j(\beta+\delta))]<1 \Rightarrow (1-j(\beta+\delta))<1 \Rightarrow -j(\beta+\delta)<0 \Rightarrow \beta+\delta<0 \Rightarrow \delta>-\beta$$

Es decir, para que el sistema sea estable, la pendiente de la curva de oferta debe superar a la pendiente de la curva de demanda. Existen tres casos de análisis posibles.

Caso 1: $\delta > -\beta$ y $\beta > 0$

Este es el caso normal en el que la curva de demanda tiene pendiente negativa. Aquí el sistema alcanza un steady state en donde el precio es positivo.

Caso 2: $\delta > -\beta$ y $\beta < 0$

Este es el caso anormal en el que la curva de demanda tiene pendiente positiva. Aquí el sistema alcanza un steady state en donde el precio es positivo.

Caso 3: $\delta = -\beta$

En este caso el steady state ***NO*** está definido dado que la división por cero no existe.

Caso 4: $\delta < -\beta$ y $\beta < 0$ **.**

Aquí los resultados no tienen sentido económico dado que el precio de equilibrio es negativo. Sin embargo, en términos matemáticos el resultado es que el sistema ***EXPLOTA*** y nunca converge a su steady state negativo.

Caso5: $\delta < -\beta$ y $\beta > 0$ **.**

Este caso no se puede dar nunca dado que suponemos que $\delta > 0$.

Ecuaciones en Diferencias No Lineales de Primer Orden: Modelo de Crecimiento de Sollow

En este caso no existe un método general que nos permita encontrar una forma reducida que describa a la relación entre y_t y y_{t+1} como función de parámetros iniciales. Es por esta razón que en este caso hay que recurrir al análisis cualitativo representado por la curva de fase. Un ejemplo clásico que podemos analizar es el modelo de Sollow.

El modelo de Sollow analiza la relación existente entre tasa de crecimiento poblacional (λ), tasa de crecimiento del stock de capital (es decir, ahorro, s) y la tecnología de que dispone la economía, la cual está representada por una función de producción Cobb-Douglas con la siguiente forma:

$$\frac{L(L^{\alpha}K^{1-\alpha})}{L} = \frac{L(L^{\alpha}K^{1-\alpha})}{L^{\alpha}L^{(1-\alpha)}} = L((\frac{K}{L})^{(1-\alpha)}) = L((k)^{(1-\alpha)}) \Rightarrow$$

$$F(L,K) = L^{\alpha}K^{1-\alpha} \Rightarrow F(L,K) = Lf(K/L,1) = Lf(k)$$

Es muy importante distinguir entre K_t y k_t. K_t es el stock de capital del que dispone esta economía en el periodo t. k_t es el stock de capital ***PER CAPITA*** del que dispone esta economía en el periodo t. La esencia del modelo de Sollow es determinar qué condiciones deben darse para que k_t alcance un valor estacionario.

La variación dinámica del stock de capital está determinada por la siguiente ecuación:

$$K_{t+1} - K_t = sF(L,K) = sL_t f(k_t) \qquad \text{A}$$

Es decir, Sollow supuso que esta economía ahorra una proporción de su producción periódica, esta es la única forma en la que puede crecer el stock de capital.

Pero recordemos además que la siguiente relación está presente:

$$k_t = K_t / L_t \Rightarrow K_t = L_t k_t \Rightarrow K_{t+1} - K_t = (L_{t+1} - L_t)k_t + L_t(k_{t+1} - k_t) \qquad \text{B}$$

Relacionando A y B tenemos que:

$$sL_t f(k_t) = (L_{t+1} - L_t)k_t + L_t(k_{t+1} - k_t)$$

Si dividimos todos los términos por L_t llegamos a lo siguiente:

$$sf(k_t) = [(L_{t+1} - L_t)/L_t]k_t + (k_{t+1} - k_t)$$

Esto implica que:

$$k_{t+1} - k_t = sf(k_t) - \lambda k_t$$

Con esta ecuación diferencial Sollow se hizo famoso y ganó el premio Nobel. Es mas, esta ecuación es la ecuación de la curva de fase del modelo. En primer lugar, el modelo tiene un steady state trivial en el cual k es cero y otro no trivial en el que k es mayor que cero. Vamos a calcular el steady state suponiendo que:

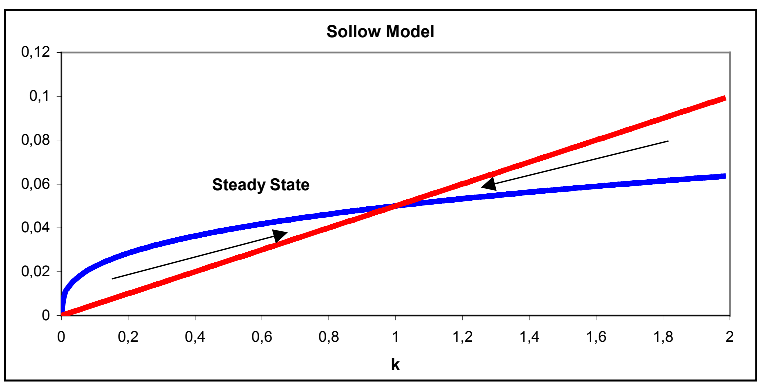
$$f(k) = k^{\alpha}$$

Vemos que si $\alpha < 1$, la función de producción exhibe rendimientos decrecientes a escala y si $\alpha > 1$, la función de producción exhibe rendimientos crecientes a escala. El steady state positivo es:

$$k_{t+1} - k_t = sf(k_t) - \lambda k_t \Rightarrow 0 = sk^{\alpha} - \lambda k \Rightarrow \lambda k = sk^{\alpha} \Rightarrow k^{1-\alpha} = s/\lambda \Rightarrow k = (s/\lambda)^{(1/(1-\alpha))}$$

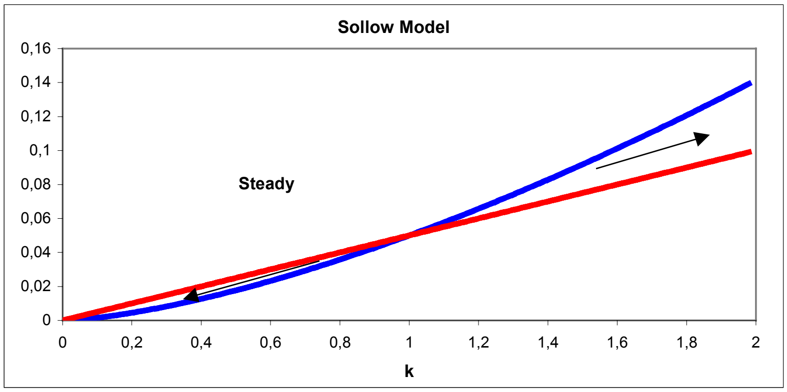
Ahora analicemos la estabilidad, vemos que:

Si $sf(k_t) > \lambda k_t \Rightarrow \Delta k_t > 0$ Si $sf(k_t) < \lambda k_t \Rightarrow \Delta k_t < 0$



Vemos que si el modelo exhibe rendimientos decrecientes a escala, su dinámica es estable y siempre converge al steady state en donde k es mayor que cero. La única razón por la que podría alcanzar el steady state en donde k es cero es si el sistema nace en el. Si observamos la dirección de las flechas, ***NO IMPORTA*** si el sistema arranca a la izquierda o a la derecha del steady state, siempre converge en el largo plazo.

Si por el contrario, el modelo exhibe rendimientos crecientes a escala, el mismo será:



Vemos aquí, que ***SI IMPORTA*** si el sistema arranca a la izquierda o la a derecha del steady state. Si arranca a la izquierda, el sistema es ***ESTABLE*** y converge a su steady state de k igual cero. Si arranca a la derecha, el sistema es ***EXPLOSIVO*** y no converge.

Como vemos, la dinámica del modelo es ***MUY RICA*** y podemos sacar las siguientes conclusiones:

Caso 1: ***Trivial, el sistema arranca en alguno de sus dos steady states: con $\bar{k}$ igual cero o $\bar{k}$ mayor que cero. En este caso, el sistema NUNCA se mueve de su steady state.***

Caso 2: ***Rendimientos decrecientes a escala. Sacando los casos triviales, no importa en dónde arranque el sistema, el mismo siempre convergerá al steady state con $\bar{k}$ mayor que cero.***

Caso 3: ***Rendimientos crecientes a escala. Sacando los casos triviales, acá sí importa en donde arranca el sistema. Si el sistema arranca con k menor a su estado estacionario positivo ($\bar{k} > 0$), el sistema converge al estado estacionario con $\bar{k}$ igual cero. Si el sistema arranca a la derecha del estado estacionario positivo, el sistema explota y no converge.***

Ecuaciones en Diferencias Lineales de Segundo Orden

Una ecuación en diferencias lineal de segundo orden tiene la siguiente forma:

$$y_{t+2} + a_1 y_{t+1} + a_2 y_t = c \qquad (1)$$

Integral Particular

Es necesario calcular el steady state del sistema:

$$\bar{y} + a_1\bar{y} + a_2\bar{y} = c \Rightarrow (1 + a_1 + a_2)\bar{y} = c \Rightarrow \bar{y} = c/(1 + a_1 + a_2) \qquad (2)$$

Por simplicidad, supongamos que $(a_1 + a_2) \neq -1$. Con lo cual, este sistema converge al steady state: $c/(1 + a_1 + a_2)$.

Función Complementaria

Como antes, este componente de la ecuación en diferencias nos indica el desvío temporal que la variable exhibe respecto a su steady state. Resolver esta parte de la ecuación en diferencias es como resolver una integral, hay que hacer un guess de la solución. Probemos con la siguiente expresión:

$$y_t = Ab^t \qquad (3)$$

Utilizando (1) y (3) simultáneamente, tenemos:

$$Ab^{t+2} + a_1 Ab^{t+1} + a_2 Ab^t = 0 \Rightarrow b^2 + a_1 b + a_2 = 0 \qquad (4)$$

Como vemos, para que el guess satisfaga la ecuación en diferencias es necesario que se satisfaga (4). Para satisfacer (4) es necesario aplicar la fórmula de la cuadratica:

$$b_1 = (-1 + \sqrt{a_1^2 - 4a_2})/2 \qquad\qquad b_2 = (-1 - \sqrt{a_1^2 - 4a_2})/2 \tag{5}$$

Dado que b_1, b_2 son raíces de un polinomio, cambiémosle la nominación:

$$\lambda_1 = b_1 \qquad\qquad \lambda_2 = b_2 \tag{6}$$

Dado que hay dos raíces en la ecuación, el guess debe expresarse como una suma:

$$y_t = A_1(\lambda_1)^t + A_2(\lambda_2)^t \tag{7}$$

Vemos que lo único desconocido en la ecuación precedente son las constantes A_1, A_2. Para encontrar la solución es entonces necesario contar con dos condiciones, por ejemplo: y_0, y_T. De esta forma, la resolución de las constantes A_1, A_2 implica armar el siguiente sistema de dos ecuaciones:

$$y_0 = A_1(\lambda_1)^0 + A_2(\lambda_2)^0 + c/(1 + a_1 + a_2) \qquad y_T = A_1(\lambda_1)^T + A_2(\lambda_2)^T + c/(1 + a_1 + a_2) \tag{8}$$

Esto implica que podemos expresarlo en álgebra matricial:

$$\begin{bmatrix} y_0 \\ y_T \end{bmatrix} = \begin{bmatrix} \lambda_1^0 & \lambda_2^0 \\ \lambda_1^T & \lambda_2^T \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} + \begin{bmatrix} c/(1 + a_1 + a_2) \\ c/(1 + a_1 + a_2) \end{bmatrix} \Rightarrow \begin{bmatrix} \lambda_1^0 & \lambda_2^0 \\ \lambda_1^T & \lambda_2^T \end{bmatrix}^{-1} \begin{bmatrix} y_0 - c/(1 + a_1 + a_2) \\ y_T - c/(1 + a_1 + a_2) \end{bmatrix} = \begin{bmatrix} \overline{A}_1 \\ \overline{A}_2 \end{bmatrix} \tag{9}$$

Una vez resuelto este sistema, el valor de las constantes A_1, A_2 es conocido, llamémoslos: $\overline{A}_1, \overline{A}_2$. De esta forma, la resolución de la ecuación en diferencias es:

$$y_t = \overline{A}_1(\lambda_1)^t + \overline{A}_2(\lambda_2)^t + c/(1 + a_1 + a_2) \tag{10}$$

<u>Programación Dinámica: Definición de Value Function, Bellman Equation y Policy Function</u>

El siguiente modelo consiste en un consumidor que está planeando consumir óptimamente durante toda su vida la cual es infinita. El individuo dispone de un stock inicial de capital k_1, y de un salario constante en cada período w. Entre un período y otro, el stock de capital produce un rendimiento de r.

$$V(k_t) = Max(\ln(c_t) + \beta V(k_{t+1})) \tag{1}$$

La restricción dinámica que enfrenta esta economía es:

$$k_{t+1} = k_t(1 + r) + w - c_t \tag{2}$$

Es decir, el stock de capital del período próximo es igual al stock de capital del período actual (k_t) mas su retorno (r), mas lo que el individuo gana en concepto de salario (w) menos lo que consume. Es decir, el ahorro en el período t de esta economía (S_t) es:

$$S_t = k_t r + w - c_t \tag{3}$$

El problema del consumidor es ahorrar óptimamente en forma dinámica. Reexprecemos (1) de la siguiente forma:

$$V(k_t) = Max(\ln(c_t) + \beta V(k_t(1 + r) + w - c_t)) \tag{4}$$

A esta ecuación se la conoce como la ***"Bellman Equation"***. En primer lugar es fundamental entender qué variable de todas estas ***CONTROLO*** en el período t. Recordemos que sólo puedo optimizar sobre aquéllas variables por las que tengo control en un punto del tiempo. En este caso, en el período t, lo único que controlo es el nivel de consumo del período t (c_t). Es muy importante reconocer que el stock de capital en el período t ***ESTA DADO*** (k_t). Es decir, en período t se determina el nivel de stock de capital aplicable para el período siguiente (k_{t+1}).

<u>Conclusión:</u> La única variable que controlo en el período t es el consumo del período t.

$V(k_t)$ es lo que se denomina ***Value Function*** (esta es una definición muy utilizada en todo lo referido a optimización intertemporal). ***La Value Function es el valor de la función objetivo evaluada en la secuencia óptima.*** La función objetivo en el período t es:

$$V(k_t) = Max_{c_t, c_{t+1}, ...}(\ln(c_t) + \beta \ln(c_{t+1}) + \beta^2 \ln(c_{t+2}) + \beta^3 \ln(c_{t+3}) + ... \Rightarrow \tag{5}$$

$$V(k_t) = Max_{c_t, c_{t+1}, ...}(\ln(c_t) + \beta[\ln(c_{t+1}) + \beta \ln(c_{t+2}) + \beta^2 \ln(c_{t+3}) + ...])$$

La función objetivo en el período $t+1$ es:

$$V(k_{t+1}) = Max_{c_{t+1}, ...}(\ln(c_{t+1}) + \beta \ln(c_{t+2}) + \beta^2 \ln(c_{t+3}) + ...) \tag{6}$$

Relacionando (5) y (6) vemos que:

$$V(k_t) = Max_{c_t, c_{t+1}, ...}(\ln(c_t) + \beta[\ln(c_{t+1}) + \beta \ln(c_{t+2}) + \beta^2 \ln(c_{t+3}) + ...]) \Rightarrow \tag{7}$$

$$V(k_t) = Max_{c_t, c_{t+1}, ...}(\ln(c_t) + \beta[V(k_{t+1})]) \Rightarrow$$

$$V(k_t) = Max_{c_t, c_{t+1}, ...}(\ln(c_t) + \beta[V(k_t(1+r) + w - c_t)])$$

Vemos que la función óptima se la define como dependiendo de la única variable de estado en este modelo (k_t). Es decir, toda nuestra optimalidad va a depender del stock inicial de capital (k_t). En realidad, el problema del consumidor es: dado el stock inicial de capital tratemos de hacer lo mejor que podamos. La condición de óptimo es:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \tag{8}$$

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}}(-1) = 0 \tag{9}$$

$$\frac{1}{c_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \Rightarrow \frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{1}{c_t \beta} \tag{10}$$

$$\frac{1}{c_{t-1}} = \beta \frac{\partial V(k_t)}{\partial k_t} \Rightarrow \frac{\partial V(k_t)}{\partial k_t} = \frac{1}{(c_{t-1})\beta}$$

Ahora derivemos a la Value Function con respecto a la variable de estado (recordemos que acá la derivada no se hace cero dado que k_t está dado y no lo controlo). A esta condición normalmente se la conoce como ***"Benveniste and Sheikman"*** condition.

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}}(1+r) \tag{11}$$

Combinando la primer ecuación en (10) con (11), tenemos:

$$\frac{\partial V(k_t)}{\partial k_t} = \beta(\frac{1}{c_t \beta})(1+r) = \frac{(1+r)}{c_t} \tag{12}$$

Combinando la segunda ecuación en (10) con (12), tenemos:

$$\frac{(1+r)}{c_t} = \frac{1}{(c_{t-1})\beta} \tag{13}$$

$$(c_{t-1})\beta(1+r) = c_t \Rightarrow c_{t+1} = c_t\beta(1+r) \tag{14}$$

Veamos lo poderosa de la expresión (14) la cual se conoce muy comúnmente como ***“la ecuación de Euler”***, ella nos dice que el intercambio óptimo entre dos períodos consecutivos depende de la relación entre el descuento intertemporal (dado por β) y el rendimiento del capital (dado por r). La expresión (14) es extremadamente importante ya que me define una relación dinámica óptima entre el consumo de un período y el consumo del período anterior. Ahora, generalicemos (14):

$$(c_1)\beta(1+r) = c_2 \tag{15}$$

$$(c_2)\beta(1+r) = c_3 \Rightarrow ((c_1)\beta(1+r))\beta(1+r) \Rightarrow c_3 = c_1\beta^2(1+r)^2$$

$$c_4 = c_1\beta^3(1+r)^3$$

Generalizando (15) tenemos:

$$c_{t+1} = c_1\beta^t(1+r)^t \Rightarrow c_t = c_1\beta^{t-1}(1+r)^{t-1} \tag{16}$$

Vemos ahora que el consumo de cualquier período futuro puede ser expresado en términos del consumo del primer período. Dada esta relación, vamos ahora a sumar el valor presente de todos los consumos:

$$\sum_{t=1}^{\infty} \frac{c_t}{(1+r)^t} \tag{17}$$

Combinando (16) y (17), tenemos:

$$\sum_{t=1}^{\infty} \frac{c_1\beta^{t-1}(1+r)^{t-1}}{(1+r)^t} = \sum_{t=1}^{\infty} \frac{c_1\beta^{t-1}}{(1+r)} = \frac{c_1}{(1+r)}[\sum_{t=1}^{\infty} \beta^{t-1}] \tag{18}$$

Recordemos que:

$$\sum_{t=1}^{\infty} \beta^{t-1} = 1+\beta+\beta^2+\beta^3+ \dots\dots\dots\dots\dots\dots\dots = \frac{1}{(1-\beta)} \tag{19}$$

Combinando (18) y (19), tenemos:

$$\sum_{t=1}^{\infty} \frac{c_1\beta^{t-1}(1+r)^{t-1}}{(1+r)^t} = \sum_{t=1}^{\infty} \frac{c_1\beta^{t-1}}{(1+r)} = \frac{c_1}{(1+r)}[\sum_{t=1}^{\infty} \beta^{t-1}] = \frac{c_1}{(1+r)}\frac{1}{(1-\beta)} \tag{20}$$

Pero recordemos que todavía no conozco a c_1. Es ahora en donde es indispensable identificar la restricción dinámica que enfrenta el consumidor. Concentrémonos en esta restricción:

$$k_{t+1} = k_t(1+r) + w - c_t \Rightarrow c_t = k_t(1+r) + w - k_{t+1} \tag{21}$$

Esto implica que:

$$\sum_{t=1}^{\infty}\frac{c_t}{(1+r)^t}=\sum_{t=1}^{\infty}\frac{k_t(1+r)}{(1+r)^t}+\sum_{t=1}^{\infty}\frac{w}{(1+r)^t}-\sum_{t=1}^{\infty}\frac{k_{t+1}}{(1+r)^t} \tag{22}$$

Concentrémonos en los dos términos que involucran al capital:

$$\sum_{t=1}^{\infty}\frac{k_t(1+r)}{(1+r)^t}=k_1+\frac{k_2}{(1+r)}+\frac{k_3}{(1+r)^2}+\frac{k_4}{(1+r)^3} \tag{23}$$

$$\sum_{t=1}^{\infty}\frac{k_{t+1}}{(1+r)^t}=\frac{k_2}{(1+r)}+\frac{k_3}{(1+r)^2}+\frac{k_4}{(1+r)^3}$$

Esto implica que:

$$\sum_{t=1}^{\infty}\frac{k_t(1+r)}{(1+r)^t}-\sum_{t=1}^{\infty}\frac{k_{t+1}}{(1+r)^t}= \tag{24}$$

$$k_1+\frac{k_2}{(1+r)}+\frac{k_3}{(1+r)^2}+\frac{k_4}{(1+r)^3}...-\frac{k_2}{(1+r)}-\frac{k_3}{(1+r)^2}-\frac{k_4}{(1+r)^3}=k_1-\frac{k_\infty}{(1+r)^{\infty-1}}$$

En primer lugar es fundamental entender qué implica el último término de (24): $k_\infty/(1+r)^{\infty-1}$. A esta condición se la denomina ***"transversality condition",*** y es una condición muy conocida. Es importante entender qué significa. La transversality condition implica una condición de optimalidad en el sentido que estamos forzando a que nuestro único recurso (es decir, en este caso, el stock de capital) sea ***CERO*** cuando termine el horizonte de optimización. La lógica de esto es muy sencilla: si el recurso tiene utilidad marginal positiva, sería subóptimo dejar algo del mismo sin usar.

Para el caso particular que estamos analizando (es decir, vida infinita), la transversality condition se cumple si al menos una de dos cosas ocurre: $k_\infty=0$ o $k_\infty/(1+r)^{\infty-1}=0$. Es decir, en un modelo de vida infinita, como el factor de descuento es "infinitamente grande", no hace falta exigir el agotamiento del recurso (es decir, $k_\infty=0$), solamente basta con exigir que su valor presente sea cero (lo cual siempre ocurre dado que estamos en un modelo de vida infinita y el factor de descuento es enorme). La situación cambia cuando el horizonte de optimización es finito (al que llamamos T). En este caso, dado que el factor de descuento ***NO*** es infinitamente grande, esto implica que $k_{T+1}/(1+r)^T>0$. En este caso, la transversality condition tiene una sola opción: $k_{T+1}=0$, lo cual también implica el agotamiento del recurso. Es importante destacar que ***ES IMPOSIBLE*** resolver este problema sin exigir una condición de transversalidad. Esta condición es tan importante como la misma ecuación de Euler.

Volviendo a nuestro problema de optimización en el infinito, como en el infinito, $k_\infty/(1+r)^{\infty-1}$ es cero, entonces:

$$\sum_{t=1}^{\infty}\frac{k_t(1+r)}{(1+r)^t}-\sum_{t=1}^{\infty}\frac{k_{t+1}}{(1+r)^t}=k_1 \tag{25}$$

Ahora analicemos los salarios en (22):

$$\sum_{t=1}^{\infty}\frac{w}{(1+r)^t}=\frac{w}{(1+r)}\sum_{t=1}^{\infty}\frac{1}{(1+r)^{t-1}} \tag{26}$$

Recordemos que:

$$\sum_{t=1}^{\infty}\frac{1}{(1+r)^{t-1}}=1+\frac{1}{(1+r)}+\frac{1}{(1+r)^2}+\frac{1}{(1+r)^3}....=\frac{1}{1-(1/(1+r))}=\frac{(1+r)}{r} \qquad (27)$$

Relacionando (26) y (27), tenemos:

$$\sum_{t=1}^{\infty}\frac{w}{(1+r)^t}=\frac{w}{(1+r)}\frac{(1+r)}{r}=\frac{w}{r} \qquad (28)$$

Ahora que tenemos cada término derivado, volvamos a (22):

$$\frac{c_1}{(1+r)}\frac{1}{(1-\beta)}=k_1+\frac{w}{r}\Rightarrow c_1=(1+r)(1-\beta)[k_1+\frac{w}{r}] \qquad (29)$$

Observamos que ya tenemos la cantidad de consumo óptima para el primer período expresada como función de la variable de estado k_1. Ahora derivemos ***TODA*** la secuencia ***OPTIMA*** de consumo. Para ello volvamos a (16):

$$c_{t+1}=c_1\beta^t(1+r)^t\Rightarrow c_t=c_1\beta^{t-1}(1+r)^{t-1} \qquad (30)$$

Combinando (29) y (30), tenemos:

$$c_t=[(1-\beta)\beta^{t-1}(1+r)^t][k_1+\frac{w}{r}] \qquad (31)$$

$$(1-\beta)=1-\frac{1}{(1+\rho)}=\frac{1+\rho-1}{(1+\rho)}=\frac{\rho}{1+\rho};\ \beta^{t-1}=\frac{1}{(1+\rho)^{t-1}}\Rightarrow$$

$$(1-\beta)\beta^{t-1}=(\frac{\rho}{1+\rho})(\frac{1}{(1+\rho)^{t-1}})=\frac{\rho}{(1+\rho)^t}\Rightarrow$$

$$c_t=[\frac{\rho}{(1+\rho)^t}(1+r)^t][k_1+\frac{w}{r}]=[\frac{(1+r)}{(1+\rho)}]^t[k_1+\frac{w}{r}]\rho$$

En la ecuación final de (31) encontramos la secuencia óptima de consumo intertemporal a la que generalmente se la denomina como ***"policy function".*** Vemos que sólo depende de la variable de estado k_1. Finalmente, si cargamos la secuencia óptima del consumo en la law of motion del capital obtendríamos la secuencia intertemporal óptima de la evolución del stock de capital. Recordemos que en (2) teníamos:

$$k_{t+1}=k_t(1+r)+w-c_t \qquad (32)$$

Entonces:

$$k_2=k_1(1+r)+w-c_1=k_1(1+r)+w-[\frac{(1+r)}{(1+\rho)}][k_1+\frac{w}{r}]\rho\Rightarrow \qquad (33)$$

$$k_2=k_1(1+r)+w-[\frac{(1+r)}{(1+\rho)}]k_1\rho-\frac{w}{r}[\frac{(1+r)}{(1+\rho)}]\rho\Rightarrow$$

$$k_2=k_1(1+r)(1-\frac{\rho}{1+\rho})+w(1-\frac{(1+r)\rho}{(1+\rho)r})\Rightarrow$$

$$1-\frac{\rho}{1+\rho}=\frac{1+\rho-\rho}{1+\rho}=\frac{1}{1+\rho};\ 1-\frac{(1+r)\rho}{(1+\rho)r}=\frac{r+\rho r-\rho-\rho r}{(1+\rho)r}=\frac{r-\rho}{(1+\rho)r}\Rightarrow$$

$$k_2=k_1(1+r)(\frac{1}{1+\rho})+w(\frac{r-\rho}{(1+\rho)r})\Rightarrow k_2=k_1a+wb$$

Vemos que a es una constante igual a:

$$a = (1+r)(\frac{1}{1+\rho}) = \frac{(1+r)}{(1+\rho)} \tag{34}$$

Vemos que b es una constante igual a:

$$b = \frac{r-\rho}{(1+\rho)r} \tag{35}$$

Volvamos entonces a (33):

$$k_2 = k_1 a + wb \tag{36}$$
$$k_3 = k_2 a + wb = (k_1 a + wb)a + wb \Rightarrow$$
$$k_3 = k_1 a^2 + wba + wb \Rightarrow k_3 = k_1 a^2 + wb(1+a) \Rightarrow$$
$$k_4 = k_3 a + wb = (k_1 a^2 + wb(1+a))a + wb = k_1 a^3 + wba + wba^2 + wb \Rightarrow$$
$$k_4 = k_1 a^3 + wb(1 + a + a^2)$$

Generalizando:

$$k_t = k_1 a^{t-1} + wb \sum_{t=1}^{t-1} a^{(t-1)} \Rightarrow \tag{37}$$
$$k_t = k_1 (\frac{(1+r)}{(1+\rho)})^{t-1} + w(\frac{r-\rho}{(1+\rho)r}) \sum_{t=1}^{t-1} (\frac{(1+r)}{(1+\rho)})^{(t-1)} \quad \forall t \geq 2$$

Vemos que esta es la secuencia óptima de acumulación de capital que se condice totalmente con la secuencia óptima del consumo. Recordemos sin embargo, que para la secuencia del stock de capital, el primer resultado de esta secuencia es conocido, dado que k_1 es la variable inicial del modelo (de ahí el $\forall t \geq 2$). Es por esta razón que la secuencia del stock de capital genera $n-1$ puntos, mientras que la secuencia del consumo genera n puntos.

***<u>Conclusión:</u>* Es importante destacar entonces cuáles son las condiciones que hay que respetar para arribar a una secuencia óptima de consumo:**

1) ***Valor inicial del stock de capital;***
2) ***Equation of motion del stock de capital;***
3) ***Ecuación de Euler para definir la relación óptima intertemporal del consumo;***
4) ***Transversality condition que exige el agotamiento del recurso al final del periodo.***

<u>Análisis de Estabilidad Dinámica</u>

Todo análisis dinámico no solamente debe concentrarse en encontrar las trayectorias temporales que optimizan la función objetivo (es decir, las policy functions), sino además en determinar si las trayectorias encontradas convergen a algún equilibrio de largo plazo. Es por esta razón que el análisis de estabilidad dinámica determina si las secuencias óptimas encontradas (es decir, policy functions) son convergentes a un equilibrio o explosivas.

En términos de análisis de estabilidad dinámica, el equilibrio de largo plazo se denomina steady state. Un steady state ocurre solamente cuando:

$$x_{t+1} = x_t \Rightarrow \Delta x = 0$$

Como vemos, un steady state se caracteriza porque la variable dinámica alcanza un equilibrio y no varía más. Esto no quiere decir que el tiempo no siga transcurriendo, lo que ocurre, es que una vez alcanzado el equilibrio dinámico, todas las variables relevantes se ***alinean*** de forma tal que el sistema no experimenta fuerzas que lo aparten del mismo. Es muy importante destacar sin embargo, que aun cuando el sistema alcanza el estado estacionario, ***TODA*** la dinámica del mismo sigue ***VIVA*** con la salvedad que las distintas fuerzas dinámicas se equilibran unas a otras.

Vamos ahora a aplicar el concepto de steady state al problema resuelto en los párrafos precedentes. Existen dos ecuaciones diferenciales claves en este problema. Una se refiere a la evolución del stock de capital y la otra a la evolución del consumo:

$$c_t\beta(1+r) = c_{t+1} \qquad k_{t+1} = k_t(1+r) + w - c_t$$

Vamos a concentrarnos en primer lugar en la ecuación diferencial del consumo. Encontrar el steady state requiere lo siguiente (recordar que $\beta = 1/(1+\rho)$):

$$c_t = c_{t+1} \Rightarrow \bar{c}\beta(1+r) = \bar{c} \Rightarrow \bar{c}((1+r)/(1+\rho)-1) = 0$$

Analicemos los siguientes casos:

<u>Caso 1</u>

Si $r = \rho$, el término $((1+r)/(1+\rho)-1)$ es cero, esto quiere decir que la condición $\bar{c}((1+r)/(1+\rho)-1) = 0$ se cumple ***AUN CUANDO*** $\bar{c}$ sea distinto de cero. Es decir, este es el ***UNICO*** caso en el que el sistema tiene dos estados estacionarios, uno con consumo positivo y el otro con consumo igual a cero.

$$\bar{c} \neq 0 \qquad \bar{c} = 0$$

Analicemos ahora el steady state del capital. Encontrar el steady state requiere lo siguiente:

$$k_{t+1} = k_t \Rightarrow 0 = \bar{k}r + w - \bar{c} \Rightarrow \bar{k} = (\bar{c} - w)/r$$

Es decir, esta condición nos dice que dado el nivel de consumo estacionario, si el stock de capital alcanza el nivel: $\bar{k} = (\bar{c} - w)/r$, el stock de capital no varía mas. Dado que en este caso el sistema puede tener dos estados estacionarios de consumo, esto implica que también existen dos estados estacionarios de capital:

$$\bar{c} \neq 0 \Rightarrow \bar{k} = (\bar{c} - w)/r \qquad \bar{c} = 0 \Rightarrow \bar{k} = -w/r$$

<u>Caso 2</u>

Si $r \neq \rho$, la única forma de que se cumpla la condición $\bar{c}((1+r)/(1+\rho)-1) = 0$ es que $\bar{c}$ sea cero. En este caso, el sistema tiene un ***SOLO*** steady state, el cual ocurre cuando el consumo es cero.

$$\bar{c} = 0$$

Analicemos ahora el steady state del capital.

$$\bar{k} = (\bar{c} - w)/r$$

Dado que en este caso el único estado estacionario de consumo se cumple cuando el mismo es cero, el estado estacionario del capital es:

$$\bar{c} = 0 \Rightarrow \bar{k} = -w/r$$

<u>Conclusión</u>: Si $r = \rho$ ***el sistema tiene dos estados estacionarios. Si*** $r \neq \rho$ ***el sistema tiene un solo estado estacionario.***

Como vemos, el sistema tiene un steady state independiente de la relación que exista entre r y ρ. Ese steady state ocurre cuando el consumo estacionario es cero y el capital estacionario es $-w/r$. Obviamente, decir que un sistema tiene steady state no quiere decir que el sistema converja al mismo. Esta convergencia dependerá de que el sistema sea estable.

Sin embargo, si se da la casualidad de que el sistema ***NACE EN ESTE STEADY STATE,*** el sistema se quedará allí para siempre, independientemente de que sea estable o inestable. Qué condición tiene que darse para que el sistema nazca estable. Lo único que tiene que darse es que el capital inicial sea:

$$k_1 = -w/r$$

Antes de seguir, y esto se aplica a todos los casos, aclaremos que el stock inicial de capital ***JAMAS*** puede ser inferior a $-w/r$. La razón es que si el stock inicial de capital es inferior, observando la policy function del consumo vemos que el mismo sería negativo, lo cual es imposible:

$$c_t = [(1-\beta)\beta^{t-1}(1+r)^t][k_1 + w/r]$$

<u>Conclusión:</u> El sistema sólo está definido para un stock inicial de capital superior a $-w/r$ **.**

Analicemos entonces, las policy functions de los distintos casos.

<u>Caso 1: Análisis de la Secuencia Intertemporal</u>

Recordemos que la policy function del consumo es:

$$c_t = [(1-\beta)\beta^{t-1}(1+r)^t][k_1 + w/r] \Rightarrow$$

$$c_t = [\frac{\rho}{(1+\rho)}(\frac{1}{(1+\rho)^{t-1}})\frac{(1+r)^t}{1}][k_1 + \frac{w}{r}] = [\frac{(1+r)^t}{(1+\rho)^t}][k_1 + \frac{w}{r}]\rho = [\frac{(1+r)^t}{(1+\rho)^t}][k_1 r + w]\frac{\rho}{r}$$

Vemos que si $r = \rho$, la policy function converge a:

$$c_t = [k_1 r + w] \Rightarrow c_t = \bar{c} \ \forall\, t$$

Vemos que esta policy function nos dice que el consumo intertemporal es ***CONSTANTE A LO LARGO DE LA VIDA DEL CONSUMIDOR***. Dada esta policy function, el estado estacionario del capital es:

$$\bar{k} = (\bar{c} - w)/r = (k_1 r + w - w)/r = k_1$$

Recordar que k_1 es el stock de capital inicial. Vemos que la solución tiene mucha intuición. El consumo de estado estacionario nos dice que el individuo se consume el rendimiento del capital mas su salario. Es decir, ***NO CONSUME EN ABSOLUTO*** el stock de capital. Es por esta razón que el valor de steady state del capital es k_1. Por otra parte, aclaremos, que el sistema ***NACE*** en steady state y ***SIEMPRE SE MANTIENE*** en steady state.

Vemos que si $k_1 = -w/r$, el consumo es cero, tal como lo describimos anteriormente:

$$c_t = [k_1 r + w] \Rightarrow c_t = [-(w/r)r + w] \Rightarrow c_t = 0 \ \forall\, t$$

Vemos entonces, que tal como lo describimos anteriormente, si el sistema nace con un stock de capital de $-w/r$, es óptimo para el consumidor (recordar que la optimalidad está definida por la policy function) no consumir nada ***NUNCA***. Pero ojo, es importante resaltar que en este caso, el sistema alcanza el steady state de consumo cero ***SOLO*** si ***NACE*** con un stock de capital de $-w/r$. Si el stock de capital es superior a $-w/r$, el sistema alcanza un steady state con consumo positivo. Es importante destacar que en este caso el sistema nace y permanece ***SIEMPRE*** en steady state.

<u>Caso 2: Análisis de la Secuencia Intertemporal</u>

En este caso se presentan dos sub-casos: $\rho > r$ y $\rho < r$.

Analicemos cuando $\rho > r$. Es importante analizar qué le ocurre a la dinámica del consumo. En realidad, la dinámica del consumo depende solamente del término:

$$(1+r)^t/(1+\rho)^t$$

Como vemos, si $\rho > r$, esto implica que ($(1+r)/(1+\rho) < 1$), lo cual implica a su vez que a medida que aumenta t, este término se aproxima a cero. La consecuencia de este resultado es que a través del tiempo ***el consumo decrece constantemente*** y se aproxima al steady state cero. En ese steady state, el steady state del capital es:

$$\bar{k} = -w/r$$

Vemos que si el sistema nace con un stock de capital de $-w/r$, obviamente se queda allí con un consumo de cero para siempre. Pero lo interesante y ***lo que lo diferencia del caso en el que*** $r = \rho$ es que aun cuando el sistema nace con un stock de capital superior a $-w/r$, el sistema ***CONVERGE*** a ese steady state naturalmente. Recordemos que en el caso en el que $r = \rho$, si el sistema nace con un stock de capital superior a $-w/r$, el sistema ***NO CONVERGE NATURALMENTE*** a ese steady state sino a uno con consumo positivo y stock de capital superior a $-w/r$.

Analicemos ahora las implicancias de un steady state en donde el capital es negativo ($-w/r$). Es decir, como el capital es negativo, el individuo permanecerá ***ETERNAMENTE ENDEUDADO.*** La pregunta es: cómo puede el individuo sostener esta deuda constante?. Recordemos que en el steady state, el individuo ya ***NO CONSUME.*** Pero recordemos, que en el steady state, el tiempo y toda su dinámica ***SIGUE*** transcurriendo, esto quiere decir que el consumidor sigue cobrando su salario w en cada periodo t. Entonces, calculemos cuál es el interés de la deuda eterna:

$$\text{Interes} = (w/r)r = w$$

Es decir, dado que el interés de la deuda eterna es exactamente igual al salario, el individuo puede sostener (y de esa forma mantener una situación de steady state) una deuda eterna pagando un interés constante financiado en su totalidad por su salario. Obvio, el costo de esto es que consumirá cero eternamente.

Analicemos cuando $\rho < r$. Como vemos, si $\rho < r$, esto implica que ($(1+r)/(1+\rho) > 1$), lo cual implica a su vez que a medida que aumenta t, este término aumenta infinitamente. La consecuencia de este resultado es que a través del tiempo ***el consumo crece constantemente*** y ***NO CONVERGE A NINGUN EQUILIBRIO ESTACIONARIO***. El único caso en el que este sistema puede quedar en steady state es ***SI NACIO ALLI (es decir, con un stock de capital inicial de*** $-w/r$).

Con lo cual, las conclusiones del modelo son:

Si $\rho = r$, los steady states son:

$$Si\ k_1 = -w/r\ \ \bar{c} = 0;\ \bar{k} = -w/r; \quad Si\ k_1 > -w/r\ \ \bar{c} > 0;\ \bar{k} = k_1$$

Es importante entender con precisión cuáles son los valores para el consumo, $\bar{c}$, y el stock de capital, $\bar{k}$, en el estado estacionario cuando $k_1 > -w/r$. Para entenderlo, concentrémonos en la law of motion del capital:

$$k_{t+1} = k_t(1+r) + w - c_t \Rightarrow k_{t+1} - k_t = k_t r + w - c_t$$

Si el sistema está en estado estacionario, sabemos que el diferencial del stock de capital es cero:

$$0 = \bar{k}r + w - \bar{c} \Rightarrow \bar{c} = \bar{k}r + w$$

Notemos que la ecuación precedente nos dice que en el estado estacionario, el individuo consume todo el rendimiento devengado del capital mas su salario, es decir, todas las fuentes de crecimiento del stock de capital. Pero no olvidemos que $\rho = r$. Esta igualdad implica que el consumo es ***SIEMPRE CONSTANTE***, es decir, el sistema ***NACE*** en el steady state sin importar cuál es el nivel inicial del stock de capital. Si el consumo es constante, también será constante el stock de capital. Con lo cual, dado que el capital no varía, su nivel de estado estacionario es el nivel inicial: k_1. De esta forma los valores del estado estacionario cuando $\rho = r$ y cuando $k_1 > -w/r$ son:

$$\bar{c} = \bar{k}_1 r + w; \ \bar{k} = k_1$$

Lo importante de este caso es que el nivel del stock de capital de estado estacionario es ***SIEMPRE*** el nivel del stock de capital con el que comienza el sistema:

$$\bar{k} = k_1$$

Obviamente, el consumo de estado estacionario es:

$$\bar{c} = \bar{k}_1 r + w$$

Esto solamente ocurre cuando $\rho = r$, es decir, ***cualquier nivel del stock de capital puede ser un estado estacionario***. Veremos que esto no ocurre en los dos siguientes casos.

Si $\rho > r$, el steady state es:

$$\bar{c} = 0; \ \bar{k} = -w/r$$

Si $\rho < r$, el sistema es ***EXPLOSIVO Y NO TIENE STEADY STATE***.

<u>Curvas de Fase</u>

El análisis anterior lo podemos hacer también con lo que se llama curvas de fase. La curva de fase para el stock de capital es:

$$\Delta k = 0 \Rightarrow k_{t+1} = k_t \Rightarrow 0 = k_t r + w - c_t \Rightarrow c_t = k_t r + w$$

En primer lugar, analicemos la forma funcional de la curva de fase:

$$\text{Si } c_t = 0 \Rightarrow 0 = k_t r + w \Rightarrow k_t = -w/r \qquad \text{Si } k_t = 0 \Rightarrow c_t = w$$

Vemos que esta curva de fase corta al eje de la variable independiente en $-w/r$, al eje de la variable dependiente en w, y tiene una pendiente de r. Qué indica una curva de fase?. La curva de fase me indica cuál debería ser el valor de la variable en cuestión (en este caso, consumo), para que dado un valor de la otra variable (en este caso el capital), el stock de capital no varíe.

$$\Delta k > 0 \Rightarrow k_t r + w - c_t > 0 \Rightarrow c_t < k_t r + w \qquad \Delta k < 0 \Rightarrow k_t r + w - c_t < 0 \Rightarrow c_t > k_t r + w$$

Es decir, siempre que estemos por debajo de la curva de fase, el stock de capital aumenta y cuando estemos por encima, el stock de capital decrece. Analicemos las curvas de fase de acuerdo a los distintos casos.

<u>Cuando</u> $\rho > r$

En este caso sabemos que el consumo decrece constantemente. La razón es la siguiente:

$$\frac{1}{c_t}(1+r) = \frac{1}{c_{t-1}\beta} \Rightarrow c_t = c_{t-1}\frac{(1+r)}{(1+\rho)} \Rightarrow c_t < c_{t-1} \text{ si } \frac{(1+r)}{(1+\rho)} < 1 \Rightarrow (1+r) < (1+\rho) \Rightarrow \rho > r$$

En segundo lugar, es también muy importante demostrar que en este caso en particular, el consumo inicial (c_1) cae por encima de la curva de fase (esto no es un detalle trivial). Recordemos que:

$$\Delta k = 0 \Rightarrow k_t r + w - c_t = 0 \Rightarrow c_t = k_t r + w \Rightarrow c_1 = k_1 r + w$$

Recordemos además, que la secuencia óptima del consumo es:

$$c_t = [(1-\beta)\beta^{t-1}(1+r)^t][k_1 + w/r] \Rightarrow c_1 = (1-\beta)(1+r)(k_1 + w/r) = ((1-\beta)(1+r)/r)(k_1 r + w)$$

Recordemos que:

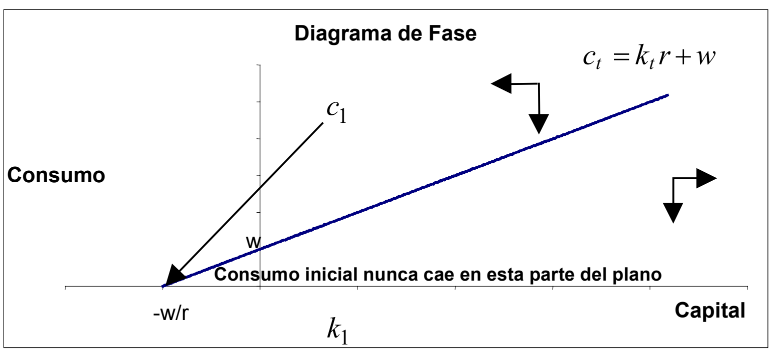
$$(1-\beta) = 1 - 1/(1+\rho) = \rho/(1+\rho)$$

Entonces:

$$c_1 = \frac{\rho(1+r)}{(1+\rho)r}(k_1 r + w) \Rightarrow c_1 > k_1 r + w \text{ si } \frac{\rho(1+r)}{(1+\rho)r} > 1 \Rightarrow \rho(1+r) > (1+\rho)r \Rightarrow$$

$$\rho + \rho r > r + \rho r \Rightarrow \rho > r$$

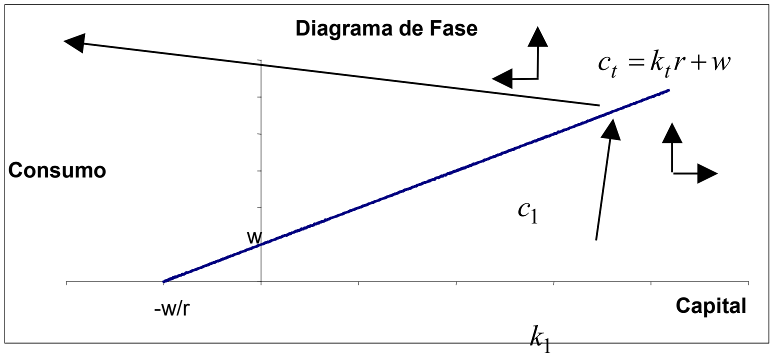
Como vemos, $c_1 > k_1 r + w$ si $\rho > r$. Entonces, cuando $\rho > r$ el consumo inicial ***SIEMPRE*** se ubica por encima de la curva de fase y decrece monotonicamente. Dado que el consumo decrece monotonicamente, las flechas verticales del gráfico (las cuales indican la dinámica del consumo) apuntan siempre hacia abajo. Las flechas horizontales del gráfico (las cuales indican la dinámica del capital) apuntan hacia la derecha si en un momento dado el consumo está por debajo de la curva de fase y para la izquierda si en un momento dado el consumo está por encima de la curva de fase***. Vemos que SOLO es relevante la parte del plano ubicado por encima de la curva de fase.***



Vemos en este caso que el sistema converge al steady state.

<u>Cuando</u> $\rho < r$

En este caso sabemos que el consumo crece constantemente. Por otra parte, usando lo que demostramos en el caso anterior, si $\rho < r$ ***el consumo inicial nunca cae por encima de la curva de fase.*** Las flechas verticales del gráfico (las cuales indican la dinámica del consumo) siempre apuntan hacia arriba dado que el consumo es monotonicamente creciente. Las flechas horizontales indican la dinámica del capital. Si el consumo cae por debajo de la curva de fase, el stock de capital es creciente, de lo contrario, es decreciente.



Vemos en este caso que el sistema es explosivo y nunca converge implicando un consumo y un endeudamiento creciente.

Cuando $\rho = r$

En este caso sabemos que el sistema siempre está en steady state. Si el capital inicial es $-w/r$, el sistema queda en un steady state con consumo igual a cero. Si el capital es superior a $-w/r$, el sistema se queda en un steady state con consumo superior a cero.

Conclusión: El comportamiento dinámico del modelo depende de la relación entre r **y** ρ **.**

Modelo Dinámico de Inflación y Desempleo: Derivación del Sendero Optimo de Output

En esta sección vamos a tratar de encontrar el sendero óptimo entre inflación y desempleo. Supongamos que tenemos un social planner que intenta ***MINIMIZAR*** una ***"social loss function"*** dada por la siguiente expresión:

$$(\bar{y} - y_t)^2 + \alpha(-\theta(\bar{y} - y_t) + \pi_t)^2 \qquad (1)$$

Antes de analizar esta función, es importante describir la dinámica del modelo. En este modelo, se supone que la tasa de inflación real en el periodo t obedece a la siguiente ley dinámica:

$$p_t = -\theta(\bar{y} - y_t) + \pi_t \qquad (2)$$

La variable $\bar{y}$ mide el nivel de ***PLENO EMPLEO*** de esta economía y y_t mide el nivel de producción de esta economía en el periodo t. El término π_t mide la ***INFLACION ESPERADA*** en el periodo t (es decir, se corresponde a las expectativas inflacionarias de los agentes). El término $-\theta(\bar{y} - y_t)$ mide el ajuste que se hace respecto a las expectativas cuando los agentes observan los datos reales de la economía. Vemos que si $\bar{y} - y_t > 0$, por ejemplo, la presión del desempleo hace que la ***VERDADERA TASA DE INFLACION***, p_t, se ajuste ***HACIA ABAJO*** respecto a la expectativa vigente, π_t.

Es muy importante tener en cuenta que coexisten dos tasas de inflación en el periodo t: a) la inflación esperada, π_t, y b) la inflación real, p_t. Para intuir mejor esta ley dinámica uno puede imaginarse que al ***INICIO*** del periodo t los agentes tienen una expectativa de inflación: π_t, la cual es ***EXOGENA*** y está dada en el periodo t. ***DURANTE*** el periodo t, los agentes deciden producir la cantidad y_t, la cual es ***ENDOGENA*** en el periodo t. De esta forma, los datos ***REALES*** de la economía se conocen, $\bar{y} - y_t$. Al ***FINAL*** del periodo t, los agentes calculan la ***VERDADERA*** tasa de inflación del periodo t la cual es igual a la expectativa inicial (π_t) mas el ***AJUSTE REAL*** ($-\theta(\bar{y} - y_t)$). Obviamente, el parámetro θ mide la importancia que el dato real tiene respecto a la expectativa inflacionaria para determinar la verdadera tasa de inflación en el periodo t. Como vemos entonces, la tasa de inflación, p_t, es la tasa de inflación esperada en t mas un ajuste real. Como vemos, es muy probable que:

$$p_t \neq \pi_t \qquad (3)$$

Lo importante ahora es entender cómo afecta a la dinámica futura del modelo el hecho que $p_t \neq \pi_t$. Si los agentes observan que se equivocaron respecto a la inflación esperada en t (es decir, $p_t \neq \pi_t$), ***AJUSTARAN SUS EXPECTATIVAS*** de inflación para el periodo siguiente, $t+1$. La ley de ajuste está determinada por la siguiente ley dinámica:

$$\pi_{t+1} = \pi_t + j(p_t - \pi_t) \qquad (4)$$

Vemos por ejemplo, que si $p_t > \pi_t$, esto generará un ajuste ***HACIA ARRIBA*** en las expectativas inflacionarias del periodo $t+1$. En este sentido, el modelo supone que los agentes exhiben ***EXPECTATIVAS ADAPTATIVAS***, la velocidad de adaptación está determinada por el parámetro j.

Si ahora combinamos las ecuaciones (2) y (4) tenemos:

$$\pi_{t+1} = \pi_t + j(p_t - \pi_t) = \pi_t + j((-\theta(\bar{y} - y_t) + \pi_t) - \pi_t) = \pi_t + j(-\theta(\bar{y} - y_t)) \Rightarrow \qquad (5)$$

$$\pi_{t+1} = \pi_t + j(-\theta(\bar{y} - y_t))$$

Es ahora en donde finalmente podemos seguir toda la dinámica del modelo. La secuencia es la siguiente:

1- En el periodo t los agentes tienen ***DADA*** una expectativa inflacionaria: π_t;
2- Los agentes toman la decisión de producir en el periodo t: y_t;
3- Los pasos (1) y (2) permiten aplicar la ecuación de inflación real, $p_t = -\theta(\bar{y} - y_t) + \pi_t$, es decir, los agentes conocen al final del periodo t la verdadera inflación, p_t;
4- En el periodo t los agentes ajustan su expectativa para el próximo periodo utilizando la ecuación $\pi_{t+1} = \pi_t + j(p_t - \pi_t)$. Al final del periodo t los agentes formaron su expectativa de inflación para el próximo periodo: π_{t+1}.

Es muy importante entender qué es exógeno en el periodo t. Lo único que es exógeno es la tasa de inflación esperada al inicio del periodo t: π_t. Las variables endogenas del periodo t son: y_t, p_t, π_{t+1}. Es muy importante tener en cuenta que la expectativa inflacionaria para el próximo periodo ($t+1$), se determina al final del periodo t.

Ahora sí podemos volver a la función de social loss en t:

$$SL_t = (\bar{y} - y_t)^2 + \alpha(-\theta(\bar{y} - y_t) + \pi_t)^2 \qquad (6)$$

Es importante entender qué mide esta función. Como vemos, la función en el periodo t está compuesta por dos elementos:

1- El primer elemento, $(\bar{y} - y_t)^2$, mide la parte de la pérdida social correspondiente al periodo t generada por ***APARTARSE*** del pleno empleo. El término se eleva al cuadrado para que estar en sobreempleo o en desempleo genere la misma pérdida. De esta forma, esta economía está en desempleo si $(\bar{y} - y_t) > 0$, y está en sobreempleo si $(\bar{y} - y_t) < 0$.

2- El segundo término, $\alpha(-\theta(\bar{y} - y_t) + \pi_t)^2$, mide la parte de la pérdida social correspondiente al periodo t generada por tener inflación o deflación. El parámetro α mide ***CUANTO*** daño le hace a esta economía tener una tasa de inflación (o deflación) dada.

Ahora que ya entendemos bien cual es la función de pérdida social, vamos a definir cuál es la función objetivo del social planner. El objetivo que tendrá el social planner es ***MINIMIZAR*** la ***SUMA INTERTEMPORAL*** de la pérdida social desde $t = 0$ hasta $t = T$. ***SUPONEMOS*** que el social planner tiene un horizonte finito de optimización determinado por T. Con lo cual, su función objetivo (es decir, la función intertemporal de social loss) es:

$$Funcion\ Objetivo: \sum_{t=0}^{T} \beta^t SL_t = \sum_{t=0}^{T} \beta^t ((\bar{y} - y_t)^2 + \alpha(-\theta(\bar{y} - y_t) + \pi_t)^2) \tag{7}$$

De esta forma, el objetivo del social planner es minimizar la ecuación precedente. La variable de control será el ingreso en cada periodo, y_t. Con lo cual, el objetivo del social planner es encontrar la secuencia óptima de producción:

$$\bar{y}_0, \bar{y}_1, \bar{y}_2, ... \bar{y}_T \tag{8}$$

En este caso, la barra sobre la variable implica óptimo. Finalmente, vamos a imponerle dos restricciones al modelo. La primera es obvia y se refiere al valor ***INICIAL*** de la variable de estado (π_0). La segunda se refiere al valor ***FINAL*** de la variable de estado (π_{T}). Es muy importante concentrarnos en esta condición final. En este problema supondremos que π_{T} será igual a un numero exógeno, es decir, por alguna razón al social planner se le dijo que π_{T} ***DEBE SIN EXCEPCION*** ser igual a una tasa dada π_T^R, en donde la letra R se refiere a que este valor está restringido ***EXOGENAMENTE***. Esto no tiene por qué ser así, bien podríamos plantear el problema como que π_{T} SEA también una ***VARIABLE*** de ***OPTIMIZACION***. Es decir, el social planner también podría determinar cuál debería ser el valor óptimo de π_{T} y de esa forma trataría a esta variable ***NO COMO UNA RESTRICCION*** sino como otra variable de control. Si este fuera el caso, el problema de optimización seria conceptualmente diferente.

Por el momento, vamos a suponer que tanto π_0 como π_T están dadas y representan restricciones. El problema de optimización es:

$$Min_{y_0,....y_T} \sum_{t=0}^{T} \beta^t ((\bar{y} - y_t)^2 + \alpha(-\theta(\bar{y} - y_t) + \pi_t)^2) \qquad Sujeto\ a: \tag{9}$$

$$\pi_{t+1} = \pi_t + j(-\theta(\bar{y} - y_t)); \qquad \pi_0 = \pi_0^R; \qquad \pi_T = \pi_T^R$$

Vemos que la variable de estado es π_t y la de control es y_t. De esta forma podemos ahora plantear la ***ecuación de Bellman***:

$$Min_{y_0 \cdots y_T} V(\pi_t) = (\bar{y} - y_t)^2 + \alpha(-\theta(\bar{y} - y_t) + \pi_t)^2 + \beta V(\pi_{t+1}) \qquad Sujeto\ a: \tag{10}$$

$$\pi_{t+1} = \pi_t + j(-\theta(\bar{y} - y_t)); \qquad \pi_0 = \pi_0^R; \qquad \pi_T = \pi_T^R$$

La condición de primer orden es:

$$\frac{\partial V(\pi_t)}{\partial y_t} = -2(\bar{y} - y_t) + 2\alpha\theta(-\theta(\bar{y} - y_t) + \pi_t) + \beta \frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} \frac{\partial \pi_{t+1}}{\partial y_t} = 0 \tag{11}$$

$$\frac{\partial V(\pi_t)}{\partial y_t} = -2(\bar{y} - y_t) + 2\alpha\theta(-\theta(\bar{y} - y_t) + \pi_t) + \beta \frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} j\theta = 0 \Rightarrow \tag{12}$$

$$\beta \frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} j\theta = 2(\bar{y} - y_t) - 2\alpha\theta(-\theta(\bar{y} - y_t) + \pi_t) \Rightarrow \tag{13}$$

$$\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} = \frac{2}{\beta j\theta}(\bar{y} - y_t) - \frac{2\alpha}{j\beta}(-\theta(\bar{y} - y_t) + \pi_t) \tag{14}$$

Distribuyendo términos:

$$\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} = \frac{2}{\beta j\theta}(\bar{y} - y_t) + \frac{2\alpha\theta}{j\beta}(\bar{y} - y_t) - \frac{2\alpha}{j\beta}\pi_t \tag{15}$$

Agrupando:

$$\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} = \frac{(2 + 2\alpha\theta^2)}{\beta j\theta}(\bar{y} - y_t) - \frac{2\alpha}{j\beta}\pi_t \tag{16}$$

$$\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} = \frac{2(1 + \alpha\theta^2)}{\beta j\theta}(\bar{y} - y_t) - \frac{2\alpha}{j\beta}\pi_t \tag{17}$$

Definamos las siguientes constantes:

$$K^1 = \frac{2(1 + \alpha\theta^2)}{\beta j\theta}; \qquad K^2 = -\frac{2\alpha}{j\beta} \tag{18}$$

De esta forma:

$$\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} = K^1(\bar{y} - y_t) + K^2\pi_t \tag{19}$$

Laggeando un periodo tenemos:

$$\frac{\partial V(\pi_t)}{\partial \pi_t} = K^1(\bar{y} - y_{t-1}) + K^2\pi_{t-1} \tag{20}$$

Ahora apliquemos la condición de ***Benveniste and Sheikman***:

$$\frac{\partial V(\pi_t)}{\partial \pi_t} = 2\alpha(-\theta(\bar{y} - y_t) + \pi_t) + \beta\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}}\frac{\partial \pi_{t+1}}{\partial \pi_t} \tag{21}$$

Dado que $\partial\pi_{t+1}/\partial\pi_t = 1$, tenemos:

$$\frac{\partial V(\pi_t)}{\partial \pi_t} = 2\alpha(-\theta(\bar{y} - y_t) + \pi_t) + \beta\frac{\partial V(\pi_{t+1})}{\partial \pi_{t+1}} \tag{22}$$

Utilicemos ahora (19) combinándolo con (22):

$$\frac{\partial V(\pi_t)}{\partial \pi_t} = 2\alpha(-\theta(\bar{y} - y_t) + \pi_t) + \beta(K^1(\bar{y} - y_t) + K^2\pi_t) \tag{23}$$

Esto implica que:

$$\frac{\partial V(\pi_t)}{\partial \pi_t} = 2\alpha(-\theta(\bar{y} - y_t) + \pi_t) + \beta K^1(\bar{y} - y_t) + \beta K^2\pi_t \tag{24}$$

Ahora combinemos (20) y (24):

$$K^1(\bar{y} - y_{t-1}) + K^2\pi_{t-1} = 2\alpha(-\theta(\bar{y} - y_t) + \pi_t) + \beta K^1(\bar{y} - y_t) + \beta K^2\pi_t \tag{25}$$

Dada la siguiente expresión: $\pi_{t+1} = \pi_t + j(-\theta(\bar{y} - y_t))$, tenemos que:

$$\frac{\pi_t - \pi_{t+1}}{j\theta} = \bar{y} - y_t \Rightarrow \frac{\pi_{t-1} - \pi_t}{j\theta} = \bar{y} - y_{t-1} \tag{26}$$

Pongamos (26) en (25):

$$K^1(\frac{\pi_{t-1} - \pi_t}{j\theta}) + K^2\pi_{t-1} = 2\alpha(-\theta(\frac{\pi_t - \pi_{t+1}}{j\theta}) + \pi_t) + \beta K^1(\frac{\pi_t - \pi_{t+1}}{j\theta}) + \beta K^2\pi_t \tag{27}$$

Distribuyendo términos:

$$\frac{K^1}{j\theta}\pi_{t-1} - \frac{K^1}{j\theta}\pi_t + K^2\pi_{t-1} = -\frac{2\alpha}{j}\pi_t + \frac{2\alpha}{j}\pi_{t+1} + 2\alpha\pi_t + \frac{\beta K^1}{j\theta}\pi_t - \frac{\beta K^1}{j\theta}\pi_{t+1} + \beta K^2\pi_t \tag{28}$$

Ahora pasemos todo al lado izquierdo de la igualdad y reagrupemos:

$$(\frac{K^1}{j\theta} + K^2)\pi_{t-1} + (-\frac{K^1}{j\theta} + \frac{2\alpha}{j} - 2\alpha - \frac{\beta K^1}{j\theta} - \beta K^2)\pi_t + (-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})\pi_{t+1} = 0 \tag{29}$$

Esto es igual a:

$$(-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})\pi_{t+1} + (-\frac{K^1}{j\theta} + \frac{2\alpha}{j} - 2\alpha - \frac{\beta K^1}{j\theta} - \beta K^2)\pi_t + (\frac{K^1}{j\theta} + K^2)\pi_{t-1} = 0 \tag{30}$$

Dividamos toda la expresión por el coeficiente asociado a π_{t+1}:

$$\pi_{t+1} + \frac{(-\frac{K^1}{j\theta} + \frac{2\alpha}{j} - 2\alpha - \frac{\beta K^1}{j\theta} - \beta K^2)}{(-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})}\pi_t + \frac{(\frac{K^1}{j\theta} + K^2)}{(-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})}\pi_{t-1} = 0 \tag{31}$$

Como vemos esto es una ecuación en diferencias de segundo orden:

$$\pi_{t+1} + a_1\pi_t + a_2\pi_{t-1} = 0\,; \quad a_1 = \frac{(-\frac{K^1}{j\theta} + \frac{2\alpha}{j} - 2\alpha - \frac{\beta K^1}{j\theta} - \beta K^2)}{(-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})}\,; \quad a_2 = \frac{(\frac{K^1}{j\theta} + K^2)}{(-\frac{2\alpha}{j} + \frac{\beta K^1}{j\theta})} \tag{32}$$

Para solucionar esta ecuación es necesario primero encontrar las dos raíces de la misma:

$$Ab^{t+2} + a_1Ab^{t+1} + a_2Ab^t = 0 \Rightarrow b^2 + a_1b + a_2 = 0 \Rightarrow \lambda_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_2}}{2}; \lambda_2 = \frac{-a_1 - \sqrt{a_1^2 - 4a_2}}{2} \tag{33}$$

Una vez conocidas las dos raíces, resolvemos el siguiente sistema de ecuaciones:

$$\pi_0 = A_1(\lambda_1)^0 + A_2(\lambda_2)^0 \tag{34}$$

$$\pi_T = A_1(\lambda_1)^T + A_2(\lambda_2)^T$$

Es muy importante detenerse en esta ecuación. Vemos que π_0 y π_T son dos restricciones que enfrenta esta problema y precisamente las usamos ahora para despejar las dos constantes. Es importante destacar que para identificar una secuencia se necesitan siempre dos puntos. En este caso los dos puntos son π_0 y π_T. Las constantes se encuentran resolviendo el siguiente sistema:

$$\begin{bmatrix} \pi_0 \\ \pi_T \end{bmatrix} = \begin{bmatrix} \lambda_1^0 & \lambda_2^0 \\ \lambda_1^T & \lambda_2^T \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \lambda_1^0 & \lambda_2^0 \\ \lambda_1^T & \lambda_2^T \end{bmatrix}^{-1} \begin{bmatrix} \pi_0 \\ \pi_T \end{bmatrix} = \begin{bmatrix} \overline{A}_1 \\ \overline{A}_2 \end{bmatrix} \tag{35}$$

De esta forma, la solución de la ley dinámica para la inflación esperada es:

$$\pi_t = \overline{A}_1(\overline{\lambda}_1)^t + \overline{A}_2(\overline{\lambda}_2)^t \tag{36}$$

Recordemos que (26) indica que:

$$\frac{\pi_t - \pi_{t+1}}{j\theta} = \overline{y} - y_t \Rightarrow y_t = \overline{y} - \frac{\pi_t}{j\theta} + \frac{\pi_{t+1}}{j\theta} \tag{37}$$

Combinando (36) y (37), tenemos:

$$y_t = \overline{y} - \frac{\overline{A}_1(\overline{\lambda}_1)^t + \overline{A}_2(\overline{\lambda}_2)^t}{j\theta} + \frac{\overline{A}_1(\overline{\lambda}_1)^{t+1} + \overline{A}_2(\overline{\lambda}_2)^{t+1}}{j\theta} = \overline{y} - \frac{\overline{A}_1(\overline{\lambda}_1)^t(1-\lambda_1) + \overline{A}_2(\overline{\lambda}_2)^t(1-\lambda_2)}{j\theta} \tag{38}$$

Con lo cual la ***policy function*** para nuestro social planner es:

$$y_t = \overline{y} - \frac{\overline{A}_1(\overline{\lambda}_1)^t(1-\lambda_1) + \overline{A}_2(\overline{\lambda}_2)^t(1-\lambda_2)}{j\theta} \tag{39}$$

Si el social planner manipula al ingreso de forma tal que en cada periodo el mismo sea el indicado por (39), la pérdida social generada por la combinación de desempleo e inflación será minimizada. La intuición es sencilla, si el social planner obedece esta ley óptima, la secuencia de eventos será la siguiente:

1- Dada una tasa inicial de inflación esperada π_0, el social planner utilizando (39) determina y_0;
2- Dados π_0 y y_0, esta economía determina la tasa de inflación p_0;
3- Dada la tasa de inflación real p_0, los agentes formarán sus expectativas de inflación esperada para el periodo siguiente: π_1.

El sistema seguirá de esta forma hasta alcanzar T. Lo importante que hay que recordar es que siguiendo la secuencia óptima definida en (39), el social planner afectará las expectativas de inflación futura óptimamente alcanzando de esta forma el óptimo social. El social planner tiene la posibilidad de ***MANIPULAR*** expectativas a traves del nivel de ingreso presente, y de esta manipulación surgirá una manipulación óptima que conduzca al óptimo social.

Finalmente, nos podemos preguntar por qué que la derivada de la expresión (11) sea igual a cero implica un mínimo y no un máximo. La razón es sencilla, ambos términos de la social loss function son convexos en la variable y_t implicando que cuando su derivada primera es cero, lo que se encuentra es un mínimo.

<u>Análisis del Trade-off entre Inflación y Desempleo</u>

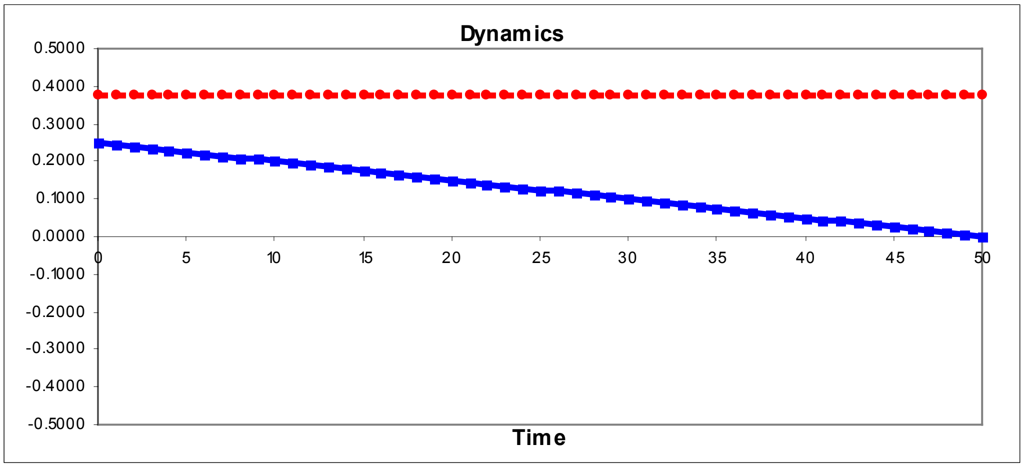
Volvamos a la social loss function:

$$SL_t = (\overline{y} - y_t)^2 + \alpha(-\theta(\overline{y} - y_t) + \pi_t)^2 \tag{40}$$

Si el social planner sólo se concentrase en reducir el desempleo, trataría de reducir el primer término de la ecuación precedente. El problema es que al reducir este término, va a simultáneamente generar mas inflación. Recordemos que de acuerdo al segundo término de la ecuación precedente, mayor inflación le ***RESTA*** bienestar a esta economía. Es por esta razón que reducir el primer término de (40) genera una simultánea suba del segundo término de (40). De ahí la existencia de ***trade-off***. Vamos a analizar algunos casos que nos permitirán ganar intuición al respecto.

Caso 1:

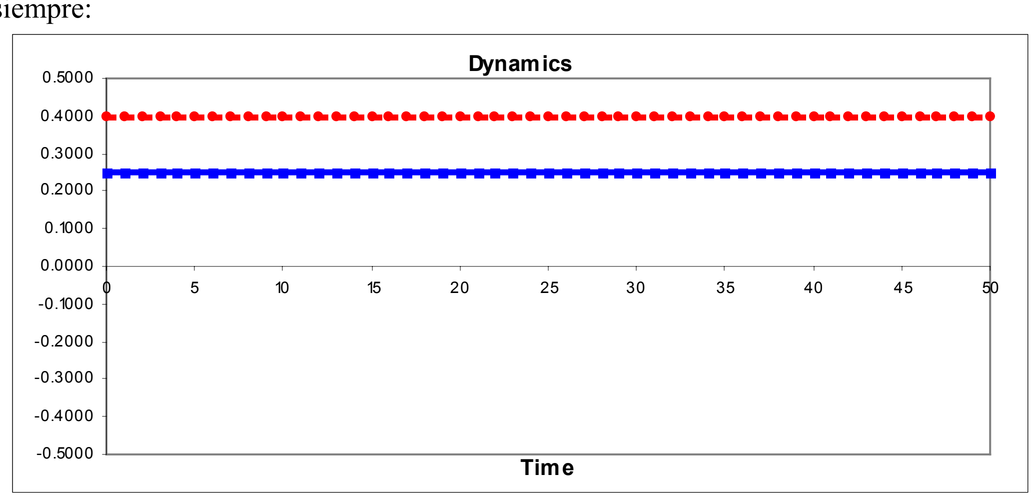
Supongamos que: $T = 50$, $\beta = 1$, $\alpha = 0$, $\bar{y} = 0.40$, $\theta = j = 0.50$, $\pi_0 = 25\%$, $\pi_T = 0\%$. Lo que nos están diciendo estos parámetros es esencialmente que a esta economía ***NO LE DUELE*** la inflación ($\alpha = 0$) y por otra parte a esta economía le ***IMPORTA MUCHO*** el futuro ($\beta = 1$). Veamos cuáles son los resultados de la secuencia óptima para el producto y la consecuente inflación esperada:



Vemos que la línea de puntos indica el nivel de ingreso mientras que la otra indica la inflación esperada. Vemos que dado que a esta economía no le duele la inflación, es óptimo estar siempre muy cerca del pleno empleo y de esta forma mantener una inflación esperada alta por mucho tiempo. En este caso es óptimo concentrarse en el primer término de la social loss function e ignorar el segundo. Sin embargo vemos que la economía no se ubica exactamente en el pleno empleo sino apenas por debajo de el. Esto se hace por una razón exclusivamente: recordemos que una de las restricciones que enfrenta el social planner es que al final del periodo cincuenta la tasa esperada de inflación debe ser $\pi_T = 0\%$. Es por esta razón que necesita un poquito de desempleo en cada periodo para ir bajando las expectativas inflacionarias desde $\pi_0 = 25\%$ hasta $\pi_T = 0\%$.

Caso 2:

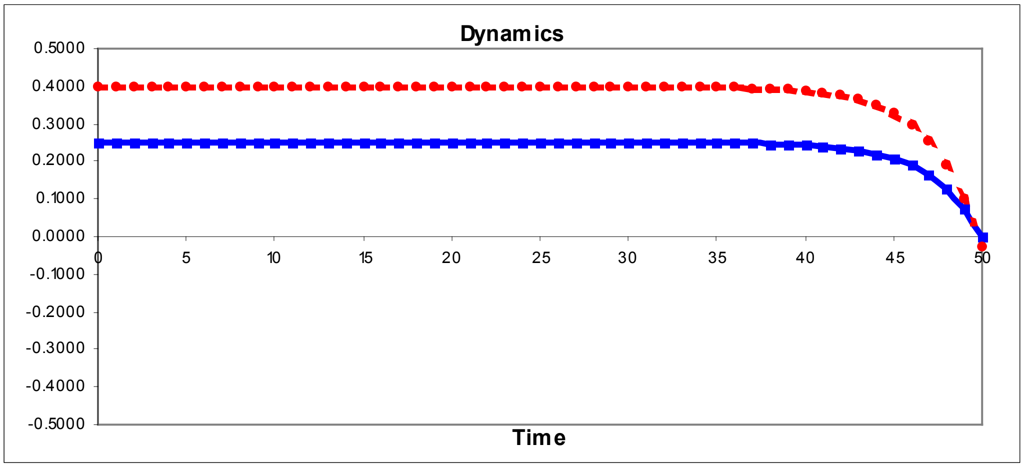
Supongamos que: $T = 50$, $\beta = 1$, $\alpha = 0$, $\bar{y} = 0.40$, $\theta = j = 0.50$, $\pi_0 = 25\%$, $\pi_T = 25\%$. Si la restricción respecto a la expectativa final hubiera sido $\pi_T = 25\%$, hubiera sido óptimo ubicar inicialmente a esta economía en el pleno empleo y dejar las expectativas de inflación constantes para siempre:



Vemos en este caso que la inflación real y la esperada son siempre iguales a 25%.

Caso 3:

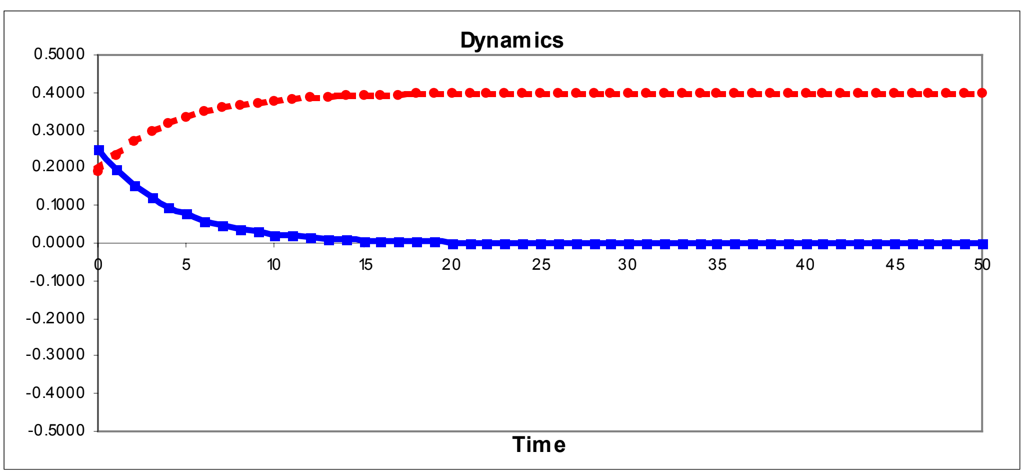
Supongamos que: $T = 50$, $\beta = 0.70$, $\alpha = 0$, $\bar{y} = 0.40$, $\theta = j = 0.50$, $\pi_0 = 25\%$, $\pi_T = 0\%$. La única diferencia respecto al ***caso 1*** es que a esta economía le importa el presente relativamente mas que el futuro ($\beta = 0.70$). Veamos cómo reacciona la dinámica a este cambio:



Vemos en este caso que el social planner se acerca al principio aun mas al pleno empleo. Recordemos que como debe converger al final de los cincuenta años a una tasa de inflación esperada del 0%, el ajuste lo hace al final. Esto es así porque si bien ese ajuste ***VA A DOLER MUCHO***, dado que a esta economía le importa relativamente mas el presente, el ajuste lo deja para el final!!!!!!!!!!!!!.

Caso 4:

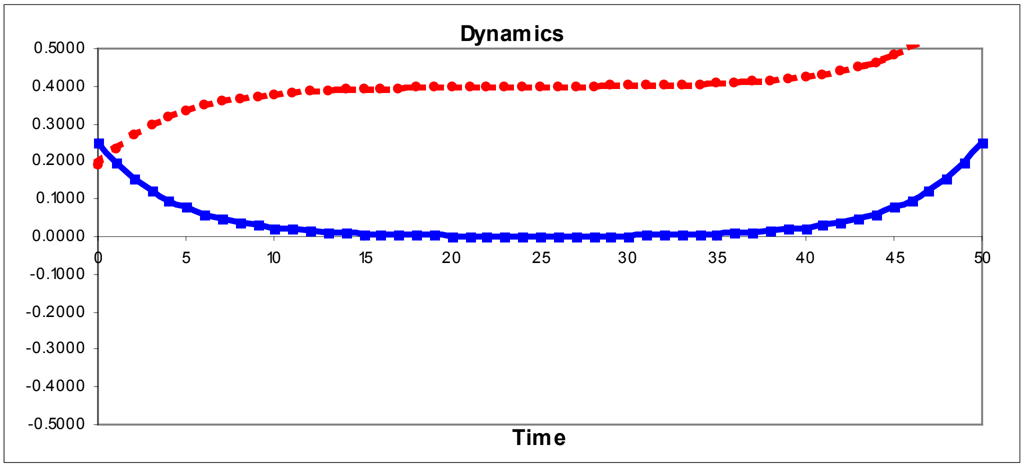
Supongamos que: $T = 50$, $\beta = 1$, $\alpha = 1$, $\bar{y} = 0.40$, $\theta = j = 0.50$, $\pi_0 = 25\%$, $\pi_T = 0\%$. La única diferencia con el ***caso 1*** es que a esta economía le ***DUELE MUCHO*** la inflación ($\alpha = 1$). Evidentemente, en este caso, el social planner deberá mantener inicialmente un nivel alto de desempleo para reducir rápidamente la inflación esperada y de esta forma reducir la inflación real. Veamos la dinámica:



Vemos que en este caso la parte curva de la línea de puntos (ingreso) representa el periodo de desempleo que tiene que sufrir esta economía para eliminar las expectativas inflacionarias. Dado que ahora a esta economía le importa la inflación, la reducción en las expectativas de inflación es rápida y brusca.

Caso 5:

Supongamos que: $T = 50$, $\beta = 1$, $\alpha = 1$, $\bar{y} = 0.40$, $\theta = j = 0.50$, $\pi_0 = 25\%$, $\pi_T = 25\%$. Las diferencias con el ***caso 1*** son que a esta economía le ***DUELE MUCHO*** la inflación ($\alpha = 1$) pero además que esta economía no tiene la obligación de bajar la tasa esperada de inflación dado que $\pi_0 = \pi_T = 25\%$. Recordemos que aunque haya que dejar permanente la expectativa de inflación al final del periodo cincuenta, a esta economía le duele mucho la inflación ($\alpha = 1$). Veamos la dinámica:



Vemos que esta dinámica es muy interesante. Lo óptimo en este caso es reducir rápidamente la inflación, dejarla baja por casi todo el periodo en cuestión, y sólo al final aumentarla rápido sobreempleando a la economía y generando de esta forma mayores expectativas de inflación y además mayor inflación real. Como conclusión, vemos que ***TODA*** la dinámica del modelo y su consecuente optimalidad dependen de las dos restricciones referidas a la variables de estado $\pi_0 = \pi_0^R$, $\pi_T = \pi_T^R$.

<u>Resolución de Sistemas Lineales de Ecuaciones en Diferencias Simultáneos</u>

<u>Sistemas Lineales Homogéneos</u>

Un sistema homogéneo para nuestro caso de análisis se caracteriza porque el steady state de ambas variables es cero. En este caso, el sistema de ecuaciones presenta la siguiente forma:

$$\begin{aligned} x_{t+1} &= a_{11}x_t + a_{12}y_t \\ y_{t+1} &= a_{21}x_t + a_{22}y_t \end{aligned} \qquad (1)$$

El steady state se resuelve de la siguiente forma:

$$\begin{aligned} \bar{x} &= a_{11}\bar{x} + a_{12}\bar{y} \\ \bar{y} &= a_{21}\bar{x} + a_{22}\bar{y} \end{aligned} \qquad (2)$$

En notación matricial, esto implica:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1-a_{22} \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (3)$$

Si el determinante asociado a la matriz es distinto de cero (lo cual descarta la posibilidad de infinitas soluciones), la única forma de resolver al sistema es cuando $\bar{x} = \bar{y} = 0$.

En contraposición, están los sistemas autónomos los cuales se caracterizan por tener un steady state para ambas variables distinto de cero. En este caso, el sistema de ecuaciones presenta la siguiente forma:

$$\begin{aligned} x_{t+1} &= b_{11} + a_{11}x_t + a_{12}y_t \\ y_{t+1} &= b_{12} + a_{21}x_t + a_{22}y_t \end{aligned} \qquad (4)$$

El steady state se resuelve así:

$$\bar{x} = b_{11} + a_{11}\bar{x} + a_{12}\bar{y} \qquad (5)$$
$$\bar{y} = b_{12} + a_{21}\bar{x} + a_{22}\bar{y}$$

En notación matricial esto implica:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1 - a_{22} \end{bmatrix} * \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} \qquad (6)$$

Entonces, la resolución de este sistema es:

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1 - a_{22} \end{bmatrix}^{-1} * \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} \qquad (7)$$

Vemos que aquí, si la matriz asociada tiene inversa, el sistema arrojará una combinación entre $\bar{x}$ y $\bar{y}$ que constituirán el estado estacionario del sistema. Vemos que la única diferencia entre ambos sistemas es que en los sistemas homogéneos la integral particular es cero mientras que en los autónomos es distinta de cero.

Vamos a analizar primero los sistemas homogéneos. En primer lugar, definamos dos conceptos básicos: eigenvalues e eigenvectors. Los eigenvalues de una matriz de 2*2 satisfacen la siguiente condición (esto es por definición):

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0 \qquad (8)$$

Esta condición genera una forma cuadratica:

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \qquad (9)$$

Esta forma cuadratica tiene dos soluciones: λ_1 y λ_2.

Los eigenvectors son aquéllos que satisfacen la siguiente condición:

$$Ae_i = \lambda_i e_i \Rightarrow \qquad (10)$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = \begin{bmatrix} \lambda_1 e_{11} \\ \lambda_1 e_{12} \end{bmatrix} \;;\; \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} = \begin{bmatrix} \lambda_2 e_{21} \\ \lambda_2 e_{22} \end{bmatrix}$$

Si extendemos la primer expresión matricial en (10), tenemos:

$$a_{11}e_{11} + a_{12}e_{12} = \lambda_1 e_{11} \Rightarrow \qquad a_{21}e_{11} + a_{22}e_{12} = \lambda_1 e_{12} \Rightarrow \qquad (11)$$
$$e_{11}(a_{11} - \lambda_1) + a_{12}e_{12} = 0 \qquad e_{11}a_{21} + e_{12}(a_{22} - \lambda_1) = 0$$

Si ahora agrupamos las dos ultimas ecuaciones de (11), nos queda que:

$$e_{11}(a_{11} - \lambda_1) + a_{12}e_{12} = 0 \qquad (12)$$

$$e_{11}a_{21} + e_{12}(a_{22} - \lambda_1) = 0$$

Esto a su vez se puede expresar como una matriz:

$$\begin{bmatrix} a_{11} - \lambda_1 & a_{12} \\ a_{21} & a_{22} - \lambda_1 \end{bmatrix} * \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (13)$$

Lo mismo ocurre si eliminamos la segunda expresión matricial en (10):

$$\begin{bmatrix} a_{11} - \lambda_2 & a_{12} \\ a_{21} & a_{22} - \lambda_2 \end{bmatrix} * \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{14}$$

Recordemos ahora un concepto básico en álgebra matricial: si el determinante de la matriz de coeficientes de un sistema lineal de ecuaciones es cero, esto implica que las ecuaciones ***SON LINEALMENTE DEPENDIENTES***. A su vez, la dependencia lineal implica que las ecuaciones ***SON REDUNDANTES***. Por ejemplo, en un sistema de dos ecuaciones, si el determinante correspondiente a la matriz de coeficientes es cero, esto implica que la segunda ecuación es redundante, por lo cual no me agrega información para resolver el sistema. Entonces, la existencia de una ecuación redundante hace que para este caso de dos ecuaciones ***ME FALTE*** una ecuación y por la tanto el sistema tenga ***INFINITAS SOLUCIONES***.

Si observamos a (13) y (14) en conjunción con (8) vemos que la matriz de coeficientes en (13) y (14) tiene por definición determinantes iguales a cero. Es por esta razón que tanto (13) como (14) ofrecen ***INFINITAS SOLUCIONES***. Esto quiere decir que hay infinitos eigenvectors que satisfacen (10). En los siguientes párrafos vamos a demostrar que esto no es en absoluto un inconveniente. Con lo cual tanto para (13) como (14) solo puedo usar una de las dos ecuaciones, resultando en la siguiente relación:

$$e_{11} = -a_{12}e_{12} / (a_{11} - \lambda_1) \qquad e_{22} = -(a_{11} - \lambda_2)e_{21} / a_{12} \tag{15}$$

Como vemos entonces, es necesario escoger arbitrariamente algún valor para e_{12} y e_{21}.

<u>Conclusión</u>: Hay infinitos eigenvectors.

Como vemos, para cada λ_i existen infinitos vectores (e_{i1}, e_{i2}) que satisfacen (13) y (14). Veamos esto con un ejemplo. Supongamos que la matriz en cuestión es:

$$\begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}$$

Las raíces características de esta matriz son dos:

$$\begin{vmatrix} 2 - \lambda_1 & 2 \\ 2 & -1-\lambda_1 \end{vmatrix} = 0 \qquad \begin{vmatrix} 2 - \lambda_2 & 2 \\ 2 & -1-\lambda_2 \end{vmatrix} = 0$$

Resolviendo la ecuación cuadrática tenemos que:

$$\lambda_1 = 3 \qquad \lambda_2 = -2$$

Ahora, aplicando (13) tenemos que:

$$\begin{bmatrix} 2-3 & 2 \\ 2 & -1-3 \end{bmatrix} * \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} * \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow e_{11} = 2e_{12}$$

$$\Rightarrow \text{Si } e_{11} = 1, e_{12} = 0.5$$

Ahora, aplicando (14) tenemos que:

$$\begin{bmatrix} 2-(-2) & 2 \\ 2 & -1-(-2) \end{bmatrix} * \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} * \begin{bmatrix} e_{21} \\ e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 4e_{21} = -2e_{22}$$

$$\Rightarrow \text{Si } e_{21} = 1, e_{22} = -2$$

Entonces, la matriz de eigenvectors es:

$$\begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0.5 & -2 \end{bmatrix}$$

Ahora bien, para qué sirve tanta álgebra? La matriz de eigenvectors tiene la siguiente propiedad:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix} * \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \Rightarrow AE = E\Lambda \Rightarrow E^{-1}AE = E^{-1}E\Lambda \tag{16}$$

De (16) surge que: (el uso de esta relación va a ser crucial próximamente):

$$E^{-1}AE = \Lambda \tag{17}$$

Ahora bien, dejemos por un instante estas relaciones y analicemos un sistema de ecuaciones en diferencias de dos variables:

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{18}$$

Recordemos que resolver un sistema de ecuaciones en diferencias implica encontrar ***SIMULTANEAMENTE*** una solución para la secuencia referida a x y y. Vemos que intentar resolver el sistema de esta forma tiene una complejidad muy importante. Expandamos el sistema:

$$x_{t+1} = a_{11}x_t + a_{12}y_t \qquad y_{t+1} = a_{21}x_t + a_{22}y_t \tag{19}$$

Vemos que (19) tiene el problema que y_t afecta a la ecuación en diferencias referida a x_{t+1}. Además, x_t afecta a la ecuación en diferencias referida a y_{t+1}. Tratar de resolver el sistema de ecuaciones en diferencias de esta forma es imposible. Es por eso que una técnica muy común es diagonalizar el sistema. Vamos a ver que a través de esta diagonalización, eliminaremos a y de la ecuación en diferencias de x y viceversa. Es ahora en donde el uso de (17) se torna fundamental. Si premultiplicamos ambos lados de (18) por la inversa de la matriz de eigenvectors tenemos:

$$E^{-1}\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = E^{-1}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_t \\ y_t \end{bmatrix} \Rightarrow E^{-1}\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = E^{-1}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} EE^{-1}\begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{20}$$

Recordemos que EE^{-1} nos da la matriz identidad y por lo tanto la igualdad no se afecta. Vemos que usando (16) tenemos que:

$$E^{-1}AE \Rightarrow E^{-1}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} E = \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{21}$$

El sistema queda expresado en variables transformadas:

$$\begin{bmatrix} \hat{x}_{t+1} \\ \hat{y}_{t+1} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} * \begin{bmatrix} \hat{x}_t \\ \hat{y}_t \end{bmatrix} \tag{22}$$

Vemos que la ***DIAGONALIZACION*** me transforma al sistema en uno de dos ecuaciones independientes:

$$\hat{x}_{t+1} = \lambda_1 \hat{x}_t \qquad \hat{y}_{t+1} = \lambda_2 \hat{y}_t \tag{23}$$

La resolución de esta ecuación en diferencias es ahora muy sencilla:

$$\hat{x}_t = \lambda_1^t c_1 \qquad \hat{y}_t = \lambda_2^t c_2 \tag{24}$$

En donde las constantes c_1 y c_2 se derivan a través de las condiciones iniciales del sistema. Retransformemos ahora nuestras variables recordando que:

$$E^{-1}\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} \hat{x}_{t+1} \\ \hat{y}_{t+1} \end{bmatrix} \Rightarrow E\begin{bmatrix} \hat{x}_{t+1} \\ \hat{y}_{t+1} \end{bmatrix} = EE^{-1}\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = \begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} \tag{25}$$

Entonces:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix} * \begin{bmatrix} \hat{x}_t \\ \hat{y}_t \end{bmatrix} = \begin{bmatrix} e_{11} & e_{21} \\ e_{12} & e_{22} \end{bmatrix} * \begin{bmatrix} \lambda_1^t c_1 \\ \lambda_2^t c_2 \end{bmatrix} \tag{26}$$

Entonces, la solución general del sistema es:

$$x_t = \lambda_1^t c_1 e_{11} + \lambda_2^t c_2 e_{21} \qquad y_t = \lambda_1^t c_1 e_{12} + \lambda_2^t c_2 e_{22} \tag{27}$$

Como vemos, ***TODO*** el comportamiento dinámico del sistema depende de las raíces λ_1 y λ_2 y de las condiciones iniciales: x_0, y_0. Estas raíces a su vez dependen ***UNICAMENTE*** de los parámetros que definen a la matriz A. Por lo tanto, el comportamiento dinámico del sistema ***UNICAMENTE*** depende de la matriz A. Recordemos que la traza de la matriz A es: $a_{11} + a_{22}$ y el determinante es $a_{11}a_{22} - a_{21}a_{12}$. Entonces:

a) Si la traza se encuentra en el rango (-2,2) y el determinante en el rango (-1,1), el sistema es globalmente estable o un saddle path. Es decir, al menos una trayectoria nos lleva al equilibrio.
b) Si la traza no se encuentra en el rango (-2,2) o el determinante no se encuentra en el rango (-1,1), el sistema es globalmente inestable o un saddle path. Es decir, cuanto mucho una trayectoria nos lleva al equilibrio.

Lo único que falta ahora es obtener dos valores iniciales para x_0 y y_0 y el sistema estaría totalmente resuelto. Vemos que, dada una matriz A, la matriz de eigenvalues y eigenvectors es totalmente conocida. Entonces, para encontrar la solución del sistema de ecuaciones en diferencias es necesario solamente encontrar valores para c_1 y c_2, los cuales se obtienen a partir de las condiciones iniciales de x_0 y y_0. Es importante destacar que aunque halla infinitas soluciones para los eigenvectors, estas infinitas soluciones ***NO*** generan infinitas secuencias. La razón es sencilla: las constantes c_1 y c_2 se ajustan de forma tal que las secuencias sean únicas. Hallemos una expresión para las constantes c_1 y c_2.

$$\begin{aligned} x_0 &= \lambda_1^0 c_1 e_{11} + \lambda_2^0 c_2 e_{21} \Rightarrow x_0 = c_1 e_{11} + c_2 e_{21} \Rightarrow c_1 = (x_0 - c_2 e_{21}) / e_{11} \\ y_0 &= \lambda_1^0 c_1 e_{12} + \lambda_2^0 c_2 e_{22} \Rightarrow y_0 = c_1 e_{12} + c_2 e_{22} \Rightarrow c_1 = (y_0 - c_2 e_{22}) / e_{12} \end{aligned} \tag{28}$$

Combinando ambas expresiones en (28) tenemos:

$$\frac{x_0 - c_2 e_{21}}{e_{11}} = \frac{y_0 - c_2 e_{22}}{e_{12}} \Rightarrow (x_0 - c_2 e_{21})e_{12} = (y_0 - c_2 e_{22})e_{11} \Rightarrow c_2 = \frac{y_0 e_{11} - x_0 e_{12}}{(e_{22}e_{11} - e_{21}e_{12})} \tag{29}$$

Una vez encontrada una expresión para c_2 podemos encontrar la equivalente para c_1. Con respecto a la estabilidad, si $abs(\lambda_1) < 1$ y $abs(\lambda_2) < 1$, el sistema es globalmente estable. Si $abs(\lambda_1) < 1$ y $abs(\lambda_2) > 1$ o $abs(\lambda_1) > 1$ y $abs(\lambda_2) < 1$, el sistema es inherentemente inestable pero exhibe un saddle path. Para "subirse al saddle path" es necesario matar la raíz inestable.

Para subirse al saddle path es necesario encontrar la ecuación del saddle path. Supongamos que la raíz inestable es λ_2 , entonces, usamos (29) y exigimos que $c_2 = 0$:

$$0 = (y_0 e_{11} - x_0 e_{12})/(e_{22}e_{11} - e_{21}e_{12}) \Rightarrow y_0 = x_0 e_{12}/e_{11} \tag{30}$$

Como vemos, la ecuación del saddle es:

$$y_0 = x_0 e_{12}/e_{11} \Rightarrow y_t = x_t e_{12}/e_{11} \tag{31}$$

Es decir, dado un valor inicial de x_0, utilizo (31) para "subirme al saddle". Una vez que encaucé al sistema en su rama estable, el sistema llega solo a su equilibrio.

<u>Sistemas Lineales Autónomos</u>

La única diferencia aquí es que la integral particular para ambas secuencias es distinta de cero y además los términos constantes del sistema ***NO*** son time-dependent. Esto implica que la solución del sistema de ecuaciones es:

$$x_t = \bar{x} + \lambda_1^t c_1 e_{11} + \lambda_2^t c_2 e_{21} \qquad y_t = \bar{y} + \lambda_1^t c_1 e_{12} + \lambda_2^t c_2 e_{22} \tag{32}$$

Vemos que la diferencia se nota en las constantes c_1 y c_2, las cuales ahora responden a la siguiente formula:

$$\begin{aligned} x_0 &= \bar{x} + \lambda_1^0 c_1 e_{11} + \lambda_2^0 c_2 e_{21} \Rightarrow x_0 = \bar{x} + c_1 e_{11} + c_2 e_{21} \Rightarrow c_1 = (x_0 - \bar{x} - c_2 e_{21})/e_{11} \\ y_0 &= \bar{y} + \lambda_1^0 c_1 e_{12} + \lambda_2^0 c_2 e_{22} \Rightarrow y_0 = \bar{y} + c_1 e_{12} + c_2 e_{22} \Rightarrow c_1 = (y_0 - \bar{y} - c_2 e_{22})/e_{12} \end{aligned} \tag{33}$$

Combinando ambas expresiones en (33) tenemos:

$$(x_0 - \bar{x} - c_2 e_{21})/e_{11} = (y_0 - \bar{y} - c_2 e_{22})/e_{12} \Rightarrow c_2 = [(y_0 - \bar{y})e_{11} - (x_0 - \bar{x})e_{12}]/(e_{22}e_{11} - e_{21}e_{12}) \tag{34}$$

Finalmente, la ecuación del saddle path ahora es:

$$c_2 = [(y_0 - \bar{y})e_{11} - (x_0 - \bar{x})e_{12}]/(e_{22}e_{11} - e_{21}e_{12}) = 0 \Rightarrow y_0 = \bar{y} + (x_0 - \bar{x})(e_{12}/e_{11}) \tag{35}$$

Es importante destacar que para la resolución de la función complementaria (la cual describe el apartamiento respecto del steady state) solo importan los términos no constantes, de ahí que para su resolución solo se considera la matriz A y no el vector B.

<u>Ejemplo: Inflación y Desempleo</u>

Un modelo clásico que analiza la relación entre la tasa de desempleo y la de inflación se basa en las siguientes tres ecuaciones en diferencias:

$$\begin{aligned} p_t &= \alpha - T - \beta U_t + h\pi_t \\ \pi_{t+1} &= \pi_t + j(p_t - \pi_t) \\ U_{t+1} &= U_t - k(m - p_{t+1}) \end{aligned} \tag{36}$$

Hay tres variables muy importantes en el modelo: p_t es la tasa real de inflación, π_t es la tasa esperada de inflación (expectativas), y U_t es la tasa de desempleo. Para empezar, este modelo exhibe enormes fallas conceptuales en el sentido que no explica por qué las tres ecuaciones en diferencias son lo que son. Recordemos por ejemplo, que para el caso del optimal growth model, la ecuación que describe la dinámica del consumo (ecuación de Euler) surge de un proceso de optimización. Es decir, la ecuación en diferencias que describe el consumo es endógena al modelo. En contraposición vemos que en este modelo las tres ecuaciones en diferencias son exógenas en el sentido que no surgen de ningún proceso de optimización inherente al modelo.

Concentrémonos en la primer ecuación. El parámetro T es la productividad del trabajo (también exógena!!!!!), el parámetro β mide la sensibilidad de la inflación real respecto a la tasa de desempleo. Recordemos que la curva de Phillips demuestra que existe una relación negativa entre tasa de desempleo e inflación. El parámetro h mide la sensibilidad de la inflación real respecto a la tasa esperada de inflación. Finalmente, el parámetro α es una constante arbitraria. Como vemos entonces, la primer ecuación nos dice que la tasa de inflación real aumenta con las expectativas de inflación y decrece con los aumentos de productividad y con la tasa de desempleo.

Vamos a la segunda ecuación, vemos aquí que estamos suponiendo un proceso de expectativas adaptativas (lo cual también está mal, dado que los agentes en general exhiben expectativas racionales). Vemos que los agentes ajustan sus expectativas de inflación de acuerdo a la diferencia entre la inflación real y esperada. La tasa a la que ajustan sus expectativas es j.

Finalmente, la tercer ecuación captura el efecto de cambios en la oferta real de dinero sobre el desempleo. El termino $m - p_{t+1}$ (el cual está expresado en tasa de crecimiento) mide la tasa de cambio de la oferta monetaria real. m es la tasa de expansión de la oferta monetaria nominal y p_{t+1} es la tasa de inflación. Obviamente, la tasa de cambio en la oferta monetaria real es la diferencia entre ambas. Finalmente el termino $-k(m - p_{t+1})$ nos indica que si la tasa de expansión de la oferta monetaria real es positiva, el desempleo disminuye. La razón es muy sencilla, incrementos de oferta monetaria real generan disminuciones de la tasa real de interés y consecuentemente, expansión económica y menor desempleo.

Si incluimos la ecuación de p_t en la ecuación de π_{t+1} y hacemos bastante álgebra, llegamos a la siguiente ecuación:

$$\begin{bmatrix} 1 & 0 \\ -kh & 1+\beta k \end{bmatrix} * \begin{bmatrix} \pi_{t+1} \\ U_{t+1} \end{bmatrix} = \begin{bmatrix} j(\alpha - T) \\ k(\alpha - T - m) \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} \pi_t \\ U_t \end{bmatrix} \tag{37}$$

Si aplicamos inversa, nos queda:

$$\begin{bmatrix} \pi_{t+1} \\ U_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -kh & 1+\beta k \end{bmatrix}^{-1} * \begin{bmatrix} j(\alpha - T) \\ k(\alpha - T - m) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -kh & 1+\beta k \end{bmatrix}^{-1} * \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} \pi_t \\ U_t \end{bmatrix} \tag{38}$$

Esto a su vez, se puede reexpresar en términos de una matriz transformada:

$$\begin{bmatrix} \pi_{t+1} \\ U_{t+1} \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} * \begin{bmatrix} \pi_t \\ U_t \end{bmatrix} \tag{39}$$

La resolución del steady state se realiza de acuerdo con (7). Es importante destacar que todo el comportamiento dinámico del modelo va a estar afectado por la matriz A la cual es:

$$\begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ -kh & 1+\beta k \end{bmatrix}^{-1} * \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{40}$$

Es decir, tanto los eigenvalues e eigenvectors que determinan la solución del sistema surgen de esta ultima matriz:

$$\pi_t = \overline{\pi} + \lambda_1^t c_1 e_{11} + \lambda_2^t c_2 e_{21} \qquad U_t = \overline{U} + \lambda_1^t c_1 e_{12} + \lambda_2^t c_2 e_{22} \tag{41}$$

<u>Ejemplo: Exchange Rate Overshooting</u>

Analicemos el modelo de exchange rate overshooting de Dornbusch. El modelo se basa en cuatro ecuaciones. El modelo ***asume perfect foresight***, lo cual significa que la tasa de inflación o devaluación que los individuos esperan coincide exactamente con la real:

$$E\left(\frac{P_{t+1}}{P_t}\right) = \frac{P_{t+1}}{P_t}; \; E\left(\frac{E_{t+1}}{E_t}\right) = \frac{E_{t+1}}{E_t} \tag{42}$$

Primero, la demanda doméstica de bienes responde a la siguiente función:

$$Y_t^d = \frac{(P^* E_t / P_t)^\delta}{(R_t^r)^\sigma}; \qquad R_t^r = \frac{R_t^N}{(1+\pi_t)} = \frac{R_t^N}{(P_{t+1}/P_t)} \tag{43}$$

En primer lugar, δ y σ son parámetros que miden elasticidades, $P^* E_t / P_t$ mide el ratio de precios relativos entre el país extranjero y el país local, R_t^r es la tasa de interés real doméstica, R_t^N es la tasa de interés nominal doméstica y π_t es la tasa de inflación doméstica. Destaquemos que suponemos que el nivel de precios internacionales es constante, de ahí que utilizamos P^* en lugar de P_t^*.

Segundo, la ecuación de inflación doméstica es:

$$\frac{P_{t+1}}{P_t} = (Y_t^d / \overline{Y})^\alpha \tag{44}$$

De nuevo, α es un parámetro que mide elasticidad. Esta ecuación nos dice que la tasa de inflación, P_{t+1} / P_t, es función del ratio entre cantidad doméstica demandada y cantidad doméstica ofrecida, $Y_t^d / \overline{Y}$. Se supone que por el lado de la oferta, la cantidad ofrecida es constante en el nivel de pleno empleo, de ahí que lleve la barra: $\overline{Y}$. Si el ratio es mayor que uno, esto implica que la demanda es mayor a la oferta, generando una presión inflacionaria. Vemos que a la oferta de bienes para simplificar se la considera rígida, de ahí que no lleva el subíndice t.

Tercero, el equilibrio en el mercado monetario requiere que la oferta monetaria real sea igual a la demanda monetaria real:

$$\frac{M}{P_t} = \frac{\overline{Y}^\theta}{(R_t^R)^\lambda} \tag{45}$$

De nuevo, θ y λ son parámetros que miden elasticidades. Vemos que la demanda monetaria real depende positivamente del ingreso real (rígido): $\overline{Y}$ y negativamente de la tasa real de interés: R_t^R. Vemos que el modelo asume una oferta monetaria nominal constante, de ahí que M carece del subíndice t.

Finalmente, el modelo incluye una relación de arbitraje entre la tasa real de interés doméstica y la tasa real de interés internacional:

$$R_t^r = R^* \frac{E_{t+1}}{E_t} \tag{46}$$

Vemos que esta ecuación simplemente nos dice que la tasa real de interés doméstica es igual a la tasa internacional mas una prima por devaluación esperada: E_{t+1} / E_t. Vemos que E_t es el tipo de cambio ***NOMINAL*** en el periodo t. Es importante destacar que suponemos que la tasa de interés internacional permanece constante en el tiempo, de ahí que utilizamos R^* en lugar de R_t^*.

Con lo cual, el modelo de Dornbusch queda constituido por las siguientes cuatro ecuaciones:

$$Y_t^d = \frac{(P^* E_t / P_t)^\delta}{(R_t^R)^\sigma} \tag{47}$$

$$\frac{P_{t+1}}{P_t} = (Y_t^d / \overline{Y})^\alpha \tag{48}$$

$$\frac{M}{P_t} = \frac{\overline{Y}^{\theta}}{(R_t^R)^{\lambda}} \qquad (49)$$

$$R_t^r = R^* \frac{E_{t+1}}{E_t} \qquad (50)$$

Si ahora sacamos logaritmos a estas cuatro ecuaciones, tenemos:

$$y_t^d = \delta(e_t + p^* - p_t) - \sigma(r_t^r) \qquad (51)$$

$$p_{t+1} - p_t = \alpha(y_t^d - \overline{y}) \qquad (52)$$

$$m - p_t = \theta\overline{y} - \lambda r_t^r \qquad (53)$$

$$r_t^r = r^* + e_{t+1} - e_t \qquad (54)$$

Vemos que las letras minúsculas indican logaritmos de las variables originales. Vemos que las dos ecuaciones ***PRINCIPALES*** del modelo son (53) y (54). (53) es una ecuación de equilibrio en el mercado monetario doméstico y (54) es una condición de equilibrio entre el mercado monetario doméstico y el internacional. Vamos ahora a combinar las cuatro ecuaciones precedentes de forma tal que nos quede un sistema de dos ecuaciones en diferencias.

Si ponemos (54) en (53), resulta:

$$m - p_t = \theta\overline{y} - \lambda(r^* + e_{t+1} - e_t) \Rightarrow \qquad (55)$$

$$m - p_t = \theta\overline{y} - \lambda r^* - \lambda(e_{t+1} - e_t) \Rightarrow \qquad (56)$$

$$\lambda(e_{t+1} - e_t) = \theta\overline{y} - \lambda r^* - m + p_t \qquad (57)$$

Ahora vamos a poner (51) en (52) pero primero recordemos que en (51) está la tasa de interés real, r_t^r. Vamos a despejar una expresión para esta tasa de interés real. Para ello, concentrémonos primero en (57) lo cual implica que:

$$(e_{t+1} - e_t) = \frac{1}{\lambda}(\theta\overline{y} - \lambda r^* - m + p_t) \qquad (58)$$

Ahora pongamos (58) en (54):

$$r_t^r = r^* + \frac{1}{\lambda}(\theta\overline{y} - \lambda r^* - m + p_t) = \frac{1}{\lambda}(\theta\overline{y} - m + p_t) \qquad (59)$$

Pongamos ahora (59) en (51):

$$y_t^d = \delta(e_t + p^* - p_t) - \frac{\sigma}{\lambda}(\theta\overline{y} - m + p_t) \tag{60}$$

Finalmente, ahora pongamos (60) en (52):

$$p_{t+1} - p_t = \alpha(\delta(e_t + p^* - p_t) - \frac{\sigma}{\lambda}(\theta\overline{y} - m + p_t) - \overline{y}) \tag{61}$$

Con lo cual, el sistema de dos ecuaciones que estamos buscando es:

$$\lambda(e_{t+1} - e_t) = \theta\overline{y} - \lambda r^* - m + p_t \tag{62}$$

$$p_{t+1} - p_t = \alpha(\delta(e_t + p^* - p_t) - \frac{\sigma}{\lambda}(\theta\overline{y} - m + p_t) - \overline{y})$$

<u>Cálculo del Steady State</u>

Si evaluamos (62) en steady state, las expresiones se simplifican a lo siguiente:

$$\lambda(\overline{e} - \overline{e}) = \theta\overline{y} - m - \lambda r^* + \overline{p} \tag{63}$$

$$\overline{p} - \overline{p} = \alpha(\delta(\overline{e} + p^* - \overline{p}) - \frac{\sigma}{\lambda}(\theta\overline{y} - m + \overline{p}) - \overline{y})$$

Concentrémonos en la primer ecuación de (63), su resolución nos permitirá encontrar el steady state value de $\overline{p}$:

$$\lambda(\overline{e} - \overline{e}) = \theta\overline{y} - m - \lambda r^* + \overline{p} \Rightarrow 0 = \theta\overline{y} - m - \lambda r^* + \overline{p} \Rightarrow \tag{64}$$
$$\overline{p} = m + \lambda r^* - \theta\overline{y}$$

Concentrémonos ahora en la segunda ecuación de (63) para derivar el steady state value de $\overline{e}$:

$$0 = \alpha(\delta(\overline{e} + p^* - \overline{p}) - \frac{\sigma}{\lambda}(\theta\overline{y} - m + \overline{p}) - \overline{y}) \Rightarrow \tag{65}$$

$$0 = \alpha\delta\overline{e} + \alpha\delta p^* - \alpha\delta\overline{p} - \frac{\alpha\sigma\theta}{\lambda}\overline{y} + \frac{\alpha\sigma}{\lambda}m - \frac{\alpha\sigma}{\lambda}\overline{p} - \alpha\overline{y} \Rightarrow \tag{66}$$

Si ahora pasamos todos los términos con excepción del primero al lado izquierdo de la igualdad, resulta:

$$-\alpha\delta p^* + \alpha\delta\overline{p} + \frac{\alpha\sigma\theta}{\lambda}\overline{y} - \frac{\alpha\sigma}{\lambda}m + \frac{\alpha\sigma}{\lambda}\overline{p} + \alpha\overline{y} = \alpha\delta\overline{e} \Rightarrow \tag{67}$$

Agrupemos términos:

$$\overline{p}(\alpha\delta + \frac{\alpha\sigma}{\lambda}) + \overline{y}(\frac{\alpha\sigma\theta}{\lambda} + \alpha) - \alpha\delta p^* - \frac{\alpha\sigma}{\lambda}m = \alpha\delta\overline{e} \tag{68}$$

Dividamos ahora a ambos términos por $\alpha\delta$:

$$\overline{p}(1 + \frac{\sigma}{\lambda\delta}) + \overline{y}(\frac{\sigma\theta}{\lambda\delta} + \frac{1}{\delta}) - p^* - \frac{\sigma}{\lambda\delta}m = \overline{e} \tag{69}$$

Reagrupando términos, resulta:

$$(\overline{p}-p^{*})+\overline{p}(\frac{\sigma}{\lambda\delta})+\overline{y}(\frac{\sigma\theta}{\lambda\delta}+\frac{1}{\delta})-\frac{\sigma}{\lambda\delta}m=\overline{e} \qquad (70)$$

Recordemos que $\overline{p}=m+\lambda r^{*}\text{-}\theta\overline{y}$:

$$(\overline{p}-p^{*})+(m+\lambda r^{*}\text{-}\theta\overline{y})(\frac{\sigma}{\lambda\delta})+\overline{y}(\frac{\sigma\theta}{\lambda\delta}+\frac{1}{\delta})-\frac{\sigma}{\lambda\delta}m=\overline{e} \qquad (71)$$

Expandiendo términos, tenemos:

$$(\overline{p}-p^{*})+m(\frac{\sigma}{\lambda\delta})+\lambda r^{*}(\frac{\sigma}{\lambda\delta})\text{-}\overline{y}(\frac{\sigma\theta}{\lambda\delta})+\overline{y}(\frac{\sigma\theta}{\lambda\delta})+\overline{y}(\frac{1}{\delta})-m(\frac{\sigma}{\lambda\delta})=\overline{e} \qquad (72)$$

Simplificando, tenemos:

$$(\overline{p}-p^{*})+\lambda r^{*}(\frac{\sigma}{\lambda\delta})+\overline{y}(\frac{1}{\delta})=\overline{e} \qquad (73)$$

Finalmente:

$$\overline{e}=(\overline{p}-p^{*})+\frac{1}{\delta}(\sigma r^{*}+\overline{y}) \qquad (74)$$

De esta forma, los steady states del sistema son:

$$\overline{p}=m+\lambda r^{*}\text{-}\theta\overline{y} \qquad\qquad \overline{e}=(\overline{p}-p^{*})+\frac{1}{\delta}(\sigma r^{*}+\overline{y}) \qquad (75)$$

<u>Análisis Dinámico</u>

Recordemos las dos ecuaciones en diferencias que describen el modelo:

$$\lambda(e_{t+1}-e_{t})=\theta\overline{y}-\lambda r^{*}-m+p_{t} \qquad (76)$$

$$p_{t+1}-p_{t}=\alpha(\delta(e_{t}+p^{*}-p_{t})-\frac{\sigma}{\lambda}(\theta\overline{y}-m+p_{t})-\overline{y})$$

Lo importante a destacar en este caso es que ambas ecuaciones ***están implícitamente*** expresadas en desviaciones de sus respectivos steady states. Obviamente, vamos a realizar un poco de álgebra para probar que (76) es un sistema expresado en desviaciones de las dos variables de sus respectivos steady states.

Empecemos con la primer ecuación de (76) la cual es la mas sencilla. Es muy importante destacar que el primer término del lado derecho de la igualdad es igual a la negativa del steady state de precios:

$$\theta\overline{y}-\lambda r^{*}-m=-\overline{p} \qquad (77)$$

Con lo cual, reemplacemos esta expresión por la negativa del steady state de precios:

$$\lambda(e_{t+1}-e_{t})=p_{t}-\overline{p} \qquad (78)$$

Por otra parte, también es importante destacar que el lado izquierdo de (78) está en desviaciones del steady state también, aunque en este caso es trivial:

$$\lambda(e_{t+1}-e_{t})=\lambda((e_{t+1}-\overline{e})-(e_{t}-\overline{e}))=\lambda(e_{t+1}-e_{t}+(\overline{e}-\overline{e}))=\lambda(e_{t+1}-e_{t}) \qquad (79)$$

Dado que $(\overline{e}-\overline{e})=0$, finalmente, la primer ecuación de (76) queda expresada así:

$$\lambda(e_{t+1}-e_{t})=p_{t}-\overline{p} \qquad (80)$$

Lo importante a destacar es que lo único que hicimos fue mostrar que esta ***ecuación siempre estuvo*** expresada en desviaciones del steady state, lo único que hicimos fue expandir el álgebra para demostrarlo.

La segunda ecuación en (76) también está expresada en desviaciones del steady state, la única diferencia es que aquí vamos a tener que hacer bastante mas álgebra para demostrarlo. Vamos a expandir el lado derecho de la igualdad:

$$p_{t+1} - p_t = \alpha(\delta e_t + \delta p^* - \delta p_t - \frac{\sigma\theta}{\lambda}\overline{y} + \frac{\sigma}{\lambda}m - \frac{\sigma}{\lambda}p_t - \overline{y}) \tag{81}$$

Agrupemos ahora a los términos e_t y p_t:

$$p_{t+1} - p_t = \alpha(\delta p^* - \frac{\sigma\theta}{\lambda}\overline{y} + \frac{\sigma}{\lambda}m - \overline{y} + \delta e_t - (\frac{\sigma}{\lambda} + \delta)p_t) \tag{82}$$

Expresar a esta ecuación en desviaciones de su steady state implica:

$$p_{t+1} - p_t = \alpha(\delta(e_t - \overline{e}) - (\frac{\sigma}{\lambda} + \delta)(p_t - \overline{p})) \tag{83}$$

Dado que dijimos que (82) está implícitamente expresada en desviaciones de su steady state, es necesario analizar qué diferencias existen entre (82) y (83). En (82) los términos que no están presentes en (83) son:

$$(\delta p^* - \frac{\sigma\theta}{\lambda}\overline{y} + \frac{\sigma}{\lambda}m - \overline{y}) = \delta p^* - \frac{\sigma\theta}{\lambda}\overline{y} + \frac{\sigma}{\lambda}m - \overline{y} \tag{84}$$

Por otra parte, en (83) los términos que no están presentes en (82) son:

$$-\delta\overline{e} + (\frac{\sigma}{\lambda} + \delta)\overline{p} \tag{85}$$

Con lo cual, para demostrar que (82) está implícitamente expresada en desviaciones de su steady state, necesitamos probar que (85) es igual a (84). Empecemos con el primer término de (85):

$$-\delta\overline{e} = -\delta((\overline{p} - p^*) + \frac{1}{\delta}(\sigma r^* + \overline{y})) \tag{86}$$

$$-\delta\overline{e} = -\delta\overline{p} + \delta p^* - \frac{\delta}{\delta}(\sigma r^* + \overline{y}) \tag{87}$$

$$-\delta\overline{e} = -\delta\overline{p} + \delta p^* - \sigma r^* - \overline{y} \tag{88}$$

Dejemos por un momento a (88), y concentrémonos en el segundo término de (85):

$$(\frac{\sigma}{\lambda} + \delta)\overline{p} = (\frac{\sigma}{\lambda} + \delta)(-\theta\overline{y} + m + \lambda r^*) \tag{89}$$

Expandiendo el álgebra, resulta:

$$(\frac{\sigma}{\lambda} + \delta)\overline{p} = (\frac{\sigma}{\lambda} + \delta)(-\theta\overline{y}) + (\frac{\sigma}{\lambda} + \delta)(m) + (\frac{\sigma}{\lambda} + \delta)(\lambda r^*) \tag{90}$$

Finalmente:

$$(\frac{\sigma}{\lambda}+\delta)\overline{p}=-\frac{\sigma\theta}{\lambda}\overline{y}-\delta\theta\overline{y}+\frac{\sigma}{\lambda}m+\delta m+\sigma r^{*}+\delta\lambda r^{*} \tag{91}$$

Con lo cual:

$$-\delta\overline{e}+(\frac{\sigma}{\lambda}+\delta)\overline{p}=-\delta\overline{p}+\delta p^{*}-\sigma r^{*}-\overline{y}-\frac{\sigma\theta}{\lambda}\overline{y}-\delta\theta\overline{y}+\frac{\sigma}{\lambda}m+\delta m+\sigma r^{*}+\delta\lambda r^{*} \tag{92}$$

Reagrupando términos y simplificando:

$$-\delta\overline{e}+(\frac{\sigma}{\lambda}+\delta)\overline{p}=\delta(-\overline{p}+(-\theta\overline{y}+m+\lambda r^{*}))+\delta p^{*}-\overline{y}-\frac{\sigma\theta}{\lambda}\overline{y}+\frac{\sigma}{\lambda}m \tag{93}$$

Vemos que el primer término del lado derecho de la igualdad es cero. Por lo tanto, (85) es igual a:

$$-\delta\overline{e}+(\frac{\sigma}{\lambda}+\delta)\overline{p}=\delta p^{*}-\frac{\sigma\theta}{\lambda}\overline{y}+\frac{\sigma}{\lambda}m-\overline{y} \tag{94}$$

Detengámonos ahora en la expresión (94). Como vemos, la misma es idéntica a la expresión contenida en (84). Esto implica que:

$$p_{t+1}-p_{t}=\alpha(\delta p^{*}-\frac{\sigma\theta}{\lambda}\overline{y}+\frac{\sigma}{\lambda}m-\overline{y}+\delta e_{t}-(\frac{\sigma}{\lambda}+\delta)p_{t})=\alpha(\delta(e_{t}-\overline{e})-(\frac{\sigma}{\lambda}+\delta)(p_{t}-\overline{p})) \tag{95}$$

De esta forma, el sistema de ecuaciones expresados en desviaciones del steady state es:

$$\lambda(e_{t+1}-e_{t})=p_{t}-\overline{p} \tag{96}$$

$$p_{t+1}-p_{t}=\alpha(\delta(e_{t}-\overline{e})-(\frac{\sigma}{\lambda}+\delta)(p_{t}-\overline{p}))$$

<u>Curva de Fase del Tipo de Cambio</u>

Por la primer ecuación de (96) sabemos que:

$$\begin{aligned}&\Delta e=e_{t+1}-e_{t}=(p_{t}-\overline{p})/\lambda\\&\Delta e=0\Rightarrow 0=(p_{t}-\overline{p})/\lambda\Rightarrow p_{t}=\overline{p}\\&\Delta e>0\Rightarrow(p_{t}-\overline{p})/\lambda>0\Rightarrow p_{t}>\overline{p}\\&\Delta e<0\Rightarrow(p_{t}-\overline{p})/\lambda<0\Rightarrow p_{t}<\overline{p}\end{aligned} \tag{97}$$

<u>Curva de Fase del Precio Doméstico</u>

Por la segunda ecuación de (96) sabemos que:

$$\Delta p=\alpha(\delta(e_{t}-\overline{e})-(\frac{\sigma}{\lambda}+\delta)(p_{t}-\overline{p})) \tag{98}$$

$$\Delta p=0\Rightarrow 0=(\delta(e_{t}-\overline{e})-(\frac{\sigma}{\lambda}+\delta)(p_{t}-\overline{p})) \tag{99}$$

$$(\frac{\sigma}{\lambda}+\delta)(p_{t}-\overline{p})=(\delta(e_{t}-\overline{e})) \tag{100}$$

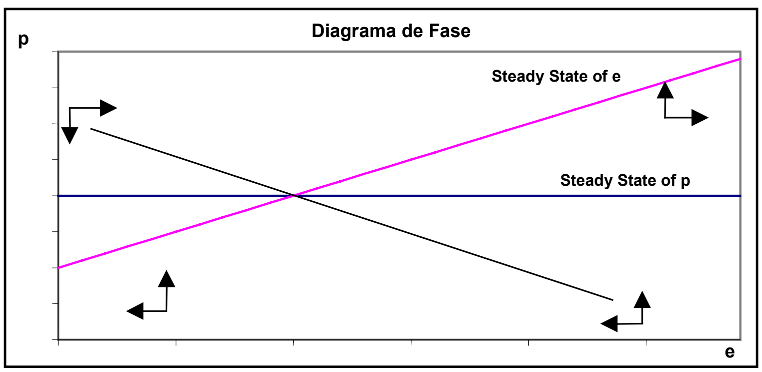
$$(\frac{\sigma}{\lambda}+\delta)p_t=(\delta(e_t-\overline{e}))+(\frac{\sigma}{\lambda}+\delta)\overline{p} \qquad (101)$$

$$p_t=\frac{\delta}{(\frac{\sigma}{\lambda}+\delta)}(e_t-\overline{e})+\overline{p} \qquad (102)$$

$$p_t=\overline{p}+\frac{\delta}{(\frac{\sigma}{\lambda}+\delta)}(e_t-\overline{e}) \qquad (103)$$

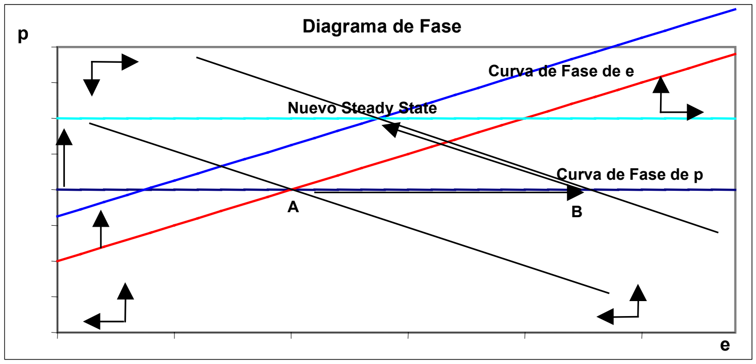
Con el mismo razonamiento:

$$\Delta p>0 \Rightarrow p_t<\overline{p}+\frac{\delta}{(\frac{\sigma}{\lambda}+\delta)}(e_t-\overline{e}) \qquad \Delta p<0 \Rightarrow p_t>\overline{p}+\frac{\delta}{(\frac{\sigma}{\lambda}+\delta)}(e_t-\overline{e}) \qquad (104)$$



Vemos que otra vez mas el sistema es ***GLOBALMENTE INESTABLE.*** Sin embargo, dado que es un punto de silla, si ***MATAMOS*** a la raíz inestable, el sistema se ***SUBE*** al saddle path y converge a su equilibrio. Fijémonos cómo partiendo de una situación de equilibrio inicial, el sistema se ajusta para converger a su nuevo equilibrio.

Supongamos que se produce un incremento en la oferta nominal de dinero. Por la primer ecuación en (75) sabemos que un incremento de m genera un valor mas alto para $\overline{p}$ (lo cual implica un desplazamiento hacia arriba de la recta horizontal que describe la curva de fase del tipo de cambio). La intuición es muy sencilla, un mayor nivel del stock nominal de dinero produce inflación superior y entonces un mayor nivel de equilibrio de precios domésticos. Además, la curva de fase de precios domésticos se traslada hacia arriba:



Vemos que un aumento de m desplaza la curva de fase del tipo de cambio (recta horizontal) hacia arriba. Además, la curva de fase del precio doméstico también se desplaza hacia arriba dado que $\overline{p}$ es la ordenada al origen de esta ecuación. Dado esto, la dinámica es la siguiente: del equilibrio inicial del tipo de cambio, dados los desplazamientos de las curvas de fase, el sistema se ***TREPA*** al saddle path (al nuevo saddle path) moviéndose ***HORIZONTALMENTE*** hacia la derecha (del punto A al B). Como vemos, este movimiento horizontal hace que ***TODO*** el ajuste lo absorba ***INICIALMENTE*** el tipo de cambio nominal, de ahí el termino ***OVERSHOOTING DEL TIPO DE CAMBIO.*** Luego de esta depreciación inicial, el sistema se acomoda en su nuevo equilibrio de largo plazo con un apreciación del tipo de cambio (vemos que después del overshooting, el tipo de cambio se aprecia un poco moviéndose hacia la izquierda) e inflación.

Es importante destacar que ***TODO*** el modelo se basa en el supuesto que los precios de los bienes domésticos son inicialmente insensibles a una suba inmediata. Esto además de ser un supuesto se condice con la realidad. Es por eso que ***TODA*** la carga del ajuste recae en el tipo de cambio nominal. La lógica del proceso de ajuste es la siguiente:

1- un aumento de oferta monetaria, dado que los precios son inicialmente rígidos a la suba, generan un aumento de la oferta monetaria real (ecuación (53)).
2- Este aumento genera una disminución en la tasa de interés nominal doméstica y consecuentemente una huida de capitales que genera una depreciación ***INMEDIATA*** del tipo de cambio nominal (ecuación (54).
3- Esta enorme depreciación del tipo de cambio nominal genera un gran incremento de demanda en favor de los bines domésticos (ecuación (51)). Este desequilibrio se soluciona con un paulatino aumento de los precios domésticos.
4- A su vez, este paulatino aumento de precios domésticos genera una disminución de la oferta monetaria real y consecuentemente una suba de la tasa de interés nominal doméstica. Esta suba en la tasa de interés genera una entrada de capitales apreciando el tipo de cambio nominal (ecuación (54)). Es por esta razón, que después del overshooting inicial, el aumento de precios permite que el tipo de cambio se ajuste paulatinamente a su nivel de equilibrio de largo plazo. Es decir, el ajuste paulatino de los precios libera al tipo de cambio nominal de soportar toda la carga del ajuste.
5- Con lo cual, un incremento de la oferta monetaria genera inicialmente una enorme depreciación del tipo de cambio nominal y luego paulatinamente y hasta converger al nuevo equilibrio, una inflación constante acompañada de una apreciación constante del tipo de cambio nominal.

Finalmente, es importante destacar lo siguiente: uno podría preguntarse por qué con el aumento inicial de la oferta monetaria, la economía se ubica ***EXACTAMENTE*** en el punto B el cual corresponde al saddle path (el cual a su vez es la única opción que tiene esta economía para converger a su nuevo equilibrio). La razón radica en el supuesto de perfect foresight el cual es un extremo del supuesto de ***rational expectations***. Decir que los agentes tienen expectativas racionales implica que los agentes conocen perfectamente todo el modelo y saben que si no van inicialmente exactamente al punto B esta economía se destruye. De ahí que en un contexto de expectativas racionales, los agentes saben exactamente que el modelo tiene un saddle y que deben treparse a el inmediatamente.

Optimal Growth Model: Equilibrio con Saddle Path

El problema que enfrenta el consumidor es el siguiente:

$$Max \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{1}$$

La restricción que enfrenta el consumidor está dada por la siguiente ecuación diferencial la cual describe la dinámica del stock de capital:

$$k_{t+1} = f(k_t) + (1-\delta)k_t - c_t \tag{2}$$

En donde $f(k_t)$ es la función de producción, c_t es el consumo en el periodo t, y δ es el factor de depreciación del stock de capital. Supongamos además que la función de utilidad es logarítmica:

$$u(c_t) = \ln(c_t) \tag{3}$$

Con lo cual, la Value Function es:

$$V(k_t) = Max(\ln(c_t) + \beta V(k_{t+1})) \tag{4}$$

Reemplazando, la restricción de capital, tenemos:

$$V(k_t) = Max(\ln(c_t) + \beta V(f(k_t) + (1-\delta)k_t - c_t) \tag{5}$$

La condición de óptimo es:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \tag{6}$$

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}}(-1) = 0 \tag{7}$$

$$\frac{1}{c_t \beta} = \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \Rightarrow \frac{1}{c_{t-1}\beta} = \frac{\partial V(k_t)}{\partial k_t} \Rightarrow \frac{\partial V(k_t)}{\partial k_t} = \frac{1}{(c_{t-1})\beta} \tag{8}$$

Ahora derivemos a la Value Function con respecto a la variable de estado (recordemos que acá la derivada no se hace cero dado que k_t está dado y no lo controlo).

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}}(f'(k_t) + (1-\delta)) \tag{9}$$

Combinando la primer ecuación en (8) con (9), tenemos:

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{\beta}{c_t \beta}(f'(k_t) + (1-\delta)) \tag{10}$$

Combinando la tercer ecuación en (8) con (10), tenemos:

$$\frac{1}{c_{t-1}\beta}=\frac{1}{c_t}(f'(k_t)+(1-\delta))\Rightarrow\frac{1}{c_t\beta}=\frac{1}{c_{t+1}}(f'(k_{t+1})+(1-\delta)) \qquad (11)$$

$$\frac{c_t\beta}{c_{t+1}}=\frac{1}{(f'(k_{t+1})+(1-\delta))} \qquad (12)$$

Veamos lo poderosa de la expresión (12), ella nos dice que el intercambio óptimo entre dos períodos consecutivos depende de la relación entre el descuento intertemporal (dado por Beta) y el rendimiento del capital neto (dado por $f'(k_{t+1})+(1-\delta)$). La expresión (12) es extremadamente importante ya que me define una relación dinámica óptima entre el consumo de un período y el consumo del período anterior y se la denomina generalmente como ***Ecuación de Euler***. Aclaremos que $f'(k_{t+1})$ mide la productividad marginal del capital.

Finalmente, el sistema queda compuesto por las siguientes dos ecuaciones diferenciales:

$$k_{t+1}=f(k_t)+(1-\delta)k_t-c_t \qquad (13)$$

$$\frac{c_t\beta}{c_{t+1}}=\frac{1}{(f'(k_{t+1})+(1-\delta))} \qquad (14)$$

<u>Calculo del Steady State</u>

$$k_{t+1}-k_t=0\Rightarrow f(\bar{k})-\delta\bar{k}-\bar{c}=0\Rightarrow\bar{c}=f(\bar{k})-\delta\bar{k} \qquad (15)$$

$$\frac{\bar{c}\beta}{\bar{c}}=\frac{1}{(f'(\bar{k})+(1-\delta))} \qquad (16)$$

Recordemos que:

$$\beta=1/(1+\rho) \qquad (17)$$

Entonces, volviendo a (16) tenemos:

$$\frac{1}{(1+\rho)}=\frac{1}{(f'(\bar{k})+(1-\delta))}\Rightarrow(f'(\bar{k})+(1-\delta))=(1+\rho)\Rightarrow f'(\bar{k})=\delta+\rho \qquad (18)$$

Con lo cual, resumiendo, las dos ecuaciones que definen el steady state son:

$$\bar{c}=f(\bar{k})-\delta\bar{k} \qquad (19)$$

$$f'(\bar{k})=\delta+\rho \qquad (20)$$

Como veremos después, vamos a usar estas dos ecuaciones en la derivación de las curvas de fase.

<u>Curvas de Fase</u>

Vamos a derivar primero la curva de fase para el capital. Recordemos que la curva de fase del capital es el locus de puntos en los que dado un valor determinado de consumo, el stock de capital permanece constante.

Curva de Fase del Capital

$$\Delta k_t = 0 \Rightarrow k_{t+1} - k_t = 0 \Rightarrow f(k_t) - \delta k_t - c_t = 0 \Rightarrow c_t = f(k_t) - \delta k_t \tag{21}$$

$$\Delta k_t > 0 \Rightarrow k_{t+1} - k_t > 0 \Rightarrow f(k_t) - \delta k_t - c_t > 0 \Rightarrow c_t < f(k_t) - \delta k_t \tag{22}$$

$$\Delta k_t < 0 \Rightarrow k_{t+1} - k_t < 0 \Rightarrow f(k_t) - \delta k_t - c_t < 0 \Rightarrow c_t > f(k_t) - \delta k_t \tag{23}$$

Curva de Fase del Consumo

Concentrémonos en la ecuación diferencial del consumo:

$$\frac{c_t \beta}{c_{t+1}} = \frac{1}{(f'(k_{t+1}) + (1-\delta))} \tag{24}$$

Entonces:

$$\frac{c_t}{c_{t+1}} = \frac{1}{\beta(f'(k_{t+1}) + (1-\delta))} \tag{25}$$

Entonces, para que Δc_t >0, debe ocurrir lo siguiente:

$$\Delta c_t > 0 \Rightarrow \frac{c_t}{c_{t+1}} < 1 \Rightarrow \frac{1}{\beta(f'(k_{t+1}) + (1-\delta))} < 1 \Rightarrow \frac{1}{(f'(k_{t+1}) + (1-\delta))} < \beta \tag{26}$$

Recordemos que $\beta = 1/(1+\rho)$, entonces:

$$\frac{1}{(f'(k_{t+1}) + (1-\delta))} < \frac{1}{(1+\rho)} \Rightarrow (1+\rho) < (f'(k_{t+1}) + (1-\delta)) \Rightarrow \rho + \delta < f'(k_{t+1}) \tag{27}$$

Recordemos que en el steady state se cumple que $\rho + \delta = f'(\bar{k})$ y además recordemos que suponemos que la función de producción es cóncava, es decir:

$$\text{Si } k_{t+1} > \bar{k} \Rightarrow f'(k_{t+1}) < f'(\bar{k}) \tag{28}$$

Dada esta explicación, volviendo a (27) tenemos:

$$\text{Si } \rho + \delta < f'(k_{t+1}) \Rightarrow f'(k_{t+1}) > \rho + \delta \Rightarrow k_{t+1} < \bar{k} \tag{29}$$

Dado que $k_{t+1} = f(k_t) + (1-\delta)k_t - c_t$, (29) puede reexpresarse como:

$$f(k_t) + (1-\delta)k_t - c_t < \bar{k} \Rightarrow c_t > f(k_t) + (1-\delta)k_t - \bar{k} \tag{30}$$

Entonces, volviendo a (26) tenemos:

$$\Delta c_t > 0 \Rightarrow c_t > f(k_t) + (1-\delta)k_t - \bar{k} \tag{31}$$

$$\Delta c_t < 0 \Rightarrow c_t < f(k_t) + (1-\delta)k_t - \bar{k} \tag{32}$$

$$\Delta c_t = 0 \Rightarrow c_t = f(k_t) + (1-\delta)k_t - \bar{k} \tag{33}$$

Análisis de la Curva de Fase del Capital

De (21) tenemos que:

$$c_t = f(k_t) - \delta k_t \tag{34}$$

Veamos el signo de la derivada:

$$dc_t / dk_t = f'(k_t) - \delta \tag{35}$$

Como vemos, si suponemos que para valores muy pequeños de k_t, $f'(k_t) > \delta$, vemos que la función tiene pendiente positiva hasta que $f'(k_t)$ empieza a ser menor que δ. Entonces:

$$\text{Si } f'(k_t) - \delta > 0 \Rightarrow \text{Funcion Creciente} \tag{36}$$

$$\text{Si } f'(k_t) - \delta < 0 \Rightarrow \text{Funcion Decreciente} \tag{37}$$

$$\text{Si } f'(k_t) - \delta = 0 \Rightarrow \text{La Funcion Alcanza su Maximo} \tag{38}$$

Además, si $k_t = 0 \Rightarrow f(k_t) - \delta k_t = 0$ (luego, vamos a usar estas conclusiones para graficar las curvas).

<u>Análisis de la Curva de Fase del Consumo</u>

De (33) tenemos que:

$$c_t = f(k_t) + (1-\delta)k_t - \bar{k} \tag{39}$$

En primer lugar, analicemos el valor de la función si $k_t = 0$:

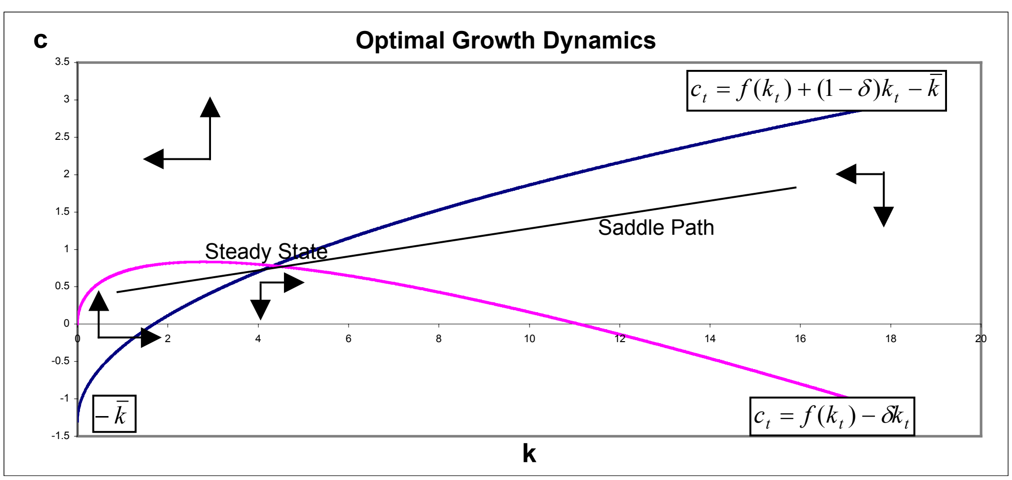
$$c_t = -\bar{k} \tag{40}$$

Ahora veamos la derivada primera:

$$dc_t / dk_t = f'(k_t) + (1-\delta) > 0 \tag{41}$$

Vemos que como esta derivada es siempre mayor que cero, la función es monotonicamente creciente y empieza en un valor negativo ($-\bar{k}$). Finalmente, vemos que cuando ambas curvas de fase se interceptan, llegamos al steady state de consumo y capital.

<u>Diagrama de Fase</u>



Como vemos, el sistema es Globalmente Inestable. Todas las secuencias con excepción de la correspondiente al saddle path nos apartan del steady state. Esto demuestra que la única opción que tenemos es ubicar a la economía en el único sendero estable: el saddle path.

Guess de la Value Function en Problemas No Lineales, No Estocásticos y con Una Sola Variable de Control

Por ejemplo, en el modelo de optimal growth, la existencia de una función de producción cóncava en la law of motion del capital introduce una no linearidad que en general impide encontrar una expresión en forma reducida para las policy functions del sistema. Sin embargo, bajo ciertos supuestos fuertes, aun en presencia de problemas no lineales es posible encontrar las policy functions en forma reducida.

La metodología que seguimos para resolver este tipo de problemas en forma reducida es a traves de una adivinanza de la Value Function. ***DADA*** esta adivinanza, se derivan las ecuaciones de Euler ***CONDICIONALES***. Una vez obtenidas estas leyes óptimas condicionales, se las vuelve a cargar en la Value Function supuesta y si se replica la forma funcional de la Value Function quiere decir que el guess fue consistente con el modelo. Es precisamente en estos casos en donde decimos que existe un contraction mapping en el sentido que el guess de la Value Function genera leyes óptimas que son consistentes con este guess y en consecuencia ***AUTORREPLICAN*** la Value Function supuesta. Una probado el hecho que el guess de la Value Function se autorreplica, se itera el proceso hasta que los parámetros de la Value Function converjan. Es decir, la Value Function se ***CONTRAE*** (***contraction mapping***) hasta que alcanza una convergencia iterativa. Vamos a analizar paso a paso esta técnica:

Paso 1: Definición del Problema de Optimización Dinámica

$$
\begin{aligned}
&Max \sum_{t=0}^{\infty} \beta^t \ln(c_t) \qquad \text{st :} \\
&k_{t+1} = k_t^{\theta} - c_t \\
&k_0 \; dado \\
&Transversality\ Condition : \lim_{t\to\infty} \beta^t u'(c_t) k_t = 0
\end{aligned} \tag{1}
$$

Como vemos, este modelo es el Optimal Growth Model con la salvedad que δ, la tasa de depreciación del capital, se la supone igual a uno, y que $f(k_t)$ se la supone igual a k_t^{θ}.

Paso 2: Cálculo del Steady State

Una de las formas de calcular el steady state es generando las ecuaciones de Euler del sistema y exigiéndoles luego constancia en el tiempo. Derivemos entonces las leyes optimas del sistema.

La ecuación de Bellman es:

$$V(k_t) = Max(\ln(c_t) + \beta V(k_{t+1})) = Max(\ln(c_t) + \beta V(k_t^{\theta} - c_t)) \tag{2}$$

La condición de primer orden respecto a la variable de control es:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \quad \frac{\partial k_{t+1}}{\partial c_t} = -1 \tag{3}$$

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} (-1) = 0 \tag{4}$$

Esta condición de primer orden implica:

$$\frac{1}{c_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \Rightarrow \frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{1}{c_t \beta}$$
$$\frac{1}{c_{t-1}} = \beta \frac{\partial V(k_t)}{\partial k_t} \Rightarrow \frac{\partial V(k_t)}{\partial k_t} = \frac{1}{c_{t-1}\beta} \qquad (5)$$

Ahora derivemos a la Value Function con respecto a la variable de estado (recordemos que acá la derivada no se hace cero dado que k_t está dado y no lo controlo). A esta condición normalmente se la conoce como ***"Benveniste and Sheikman"*** condition:

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t} \qquad \frac{\partial k_{t+1}}{\partial k_t} = \theta k_t^{(\theta-1)} \Rightarrow \qquad (6)$$
$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \theta k_t^{(\theta-1)}$$

La primer ecuación de (5) nos da una expresión para $\partial V(k_{t+1}) / \partial k_{t+1}$. Pongamos esta expresión en la última ecuación de (6):

$$\frac{\partial V(k_t)}{\partial k_t} = \beta (\frac{1}{c_t \beta}) \theta k_t^{(\theta-1)} = \frac{1}{c_t} \theta k_t^{(\theta-1)} \qquad (7)$$

Vemos que (7) nos da una expresión para $\partial V(k_t) / \partial k_t$. Vemos también que la última ecuación de (5) nos da una expresión equivalente para $\partial V(k_t) / \partial k_t$. Si relacionamos a ambas, resulta la ecuación de Euler del sistema:

$$\frac{1}{c_{t-1}\beta} = \frac{1}{c_t} \theta k_t^{(\theta-1)} \Rightarrow c_t = c_{t-1} \beta \theta k_t^{(\theta-1)} \text{ \textit{Euler Equation}} \qquad (8)$$

Como contrapartida de esta ley óptima, la ley óptima para la acumulación de capital es:

$$k_{t+1} = k_t^{\theta} - c_{t-1} \beta \theta k_t^{(\theta-1)} \text{ \textit{Euler Equation}} \qquad (9)$$

Si evaluamos la última ecuación en (8) en staedy state, resulta:

$$\bar{c} = \bar{c} \beta \theta \bar{k}^{(\theta-1)} \Rightarrow 1 = \beta \theta \bar{k}^{(\theta-1)} \Rightarrow \bar{k} = (\frac{1}{\beta\theta})^{1/(\theta-1)} \text{ \textit{Steady State}} \qquad (10)$$

Dado que ya tenemos el steady state del capital, calculemos el del consumo utilizando la law of motion del capital evaluada en el steady state:

$$k_t^{\theta} = k_t^{\theta} - c_t \Rightarrow c_t = k_t^{\theta} - k_{t+1} \Rightarrow \qquad (11)$$
$$\bar{c} = \bar{k}^{\theta} - \bar{k} \Rightarrow \bar{c} = ((\frac{1}{\beta\theta})^{1/(\theta-1)})^{\theta} - \bar{k} \Rightarrow \bar{c} = (\frac{1}{\beta\theta})^{\theta/(\theta-1)} - (\frac{1}{\beta\theta})^{1/(\theta-1)} \text{ \textit{Steady State}}$$

Con lo cual, el steady state del sistema es:

$$\bar{k} = (\frac{1}{\beta\theta})^{1/(\theta-1)} \qquad \bar{c} = (\frac{1}{\beta\theta})^{\theta/(\theta-1)} - (\frac{1}{\beta\theta})^{1/(\theta-1)} \qquad (12)$$

Paso 3: Guess de la Value Function

SUPONGAMOS que la Value Function tiene la siguiente forma funcional:

$$V(k_t) = E + F \ln(k_t) \Rightarrow V(k_{t+1}) = E + F \ln(k_{t+1}) \qquad (13)$$

Recordemos además que la Value Function alternativamente se define como:

$$V(k_t) = Max\ln(c_t) + \beta V(k_{t+1}) \qquad (14)$$

Combinando (13) y (14) obtenemos:

$$V(k_t) = Max\ln(c_t) + \beta(E + F\ln(k_{t+1})) \qquad (15)$$

Finalmente, reemplazando k_{t+1} por su law of motion, obtenemos:

$$V(k_t) = Max\ln(c_t) + \beta(E + F\ln(k_t^{\theta} - c_t)) \qquad (16)$$

De ahora en adelante, obtendremos como siempre las leyes óptimas del sistema pero con el ***DETALLE*** de que ***TODAS*** las condiciones de optimalidad ***SERAN CONDICIONALES*** en la Value Function que supusimos. De esta forma, las Euler Equations del sistema serán óptimas en la medida que el guess de la Value Function haya sido consistente con la ***ESENCIA*** del modelo.

Paso 4: Cálculo de la Condición de Primer Orden y Leyes Optimas Condicionales

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta\frac{\partial V(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial c_t} = 0 \quad \frac{\partial k_{t+1}}{\partial c_t} = -1 \qquad (17)$$

Ahora, dado el guess de la Value Function, contamos con un dato adicional que antes desconocíamos. Este dato adicional está representado por el hecho que ahora sabemos qué es $\partial V(k_{t+1}) / \partial k_{t+1}$:

$$\frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{\partial(E + F\ln(k_t^{\theta} - c_t))}{\partial k_{t+1}} = \frac{F}{(k_t^{\theta} - c_t)} \qquad (18)$$

Poniendo (18) en (17) tenemos:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta\frac{F}{(k_t^{\theta} - c_t)}(-1) = 0 \qquad (19)$$

Esto a su vez implica:

$$\frac{1}{c_t} = \beta\frac{F}{(k_t^{\theta} - c_t)} \Rightarrow (k_t^{\theta} - c_t) = c_t\beta F \Rightarrow c_t(1 + \beta F) = k_t^{\theta} \Rightarrow \qquad (20)$$

$$c_t = \frac{k_t^{\theta}}{(1 + \beta F)}$$

Es muy importante detenernos en la expresión (20). Vemos que esta expresión es la policy function del consumo ***CONDICIONAL*** en el guess de la Value Function. Si el guess de la Value Function fuera erróneo, esta ley óptima no tendría nada que ver con la verdadera ley óptima del sistema. De la misma forma, la ley óptima para la acumulación de capital es:

$$k_{t+1} = k_t^{\theta} - c_t \Rightarrow k_{t+1} = k_t^{\theta} - \frac{k_t^{\theta}}{(1 + \beta F)} \Rightarrow \qquad (21)$$

$$k_{t+1} = \frac{k_t^{\theta} + \beta F k_t^{\theta} - k_t^{\theta}}{(1 + \beta F)} \Rightarrow k_{t+1} = (\frac{\beta F}{(1 + \beta F)})k_t^{\theta}$$

Muy bien, las dos leyes optimas ***CONDICIONALES*** del sistema son:

$$c_t = \frac{k_t^{\theta}}{(1 + \beta F)} \qquad k_{t+1} = (\frac{\beta F}{(1 + \beta F)})k_t^{\theta} \qquad (22)$$

Paso 5: Determinación de la Replicación Funcional de la Value Function

Si el guess de la Value Function fue correcto, si cargásemos las leyes optimas condicionales descriptas en (22) en la Value Function supuesta, la misma debería ***AUTORREPLICARSE***. Vamos a determinar si esto en verdad ocurre:

$$V(k_t) = \ln c_t + \beta V(k_{t+1}) \Rightarrow V(k_t) = \ln(\frac{k_t^{\theta}}{(1+\beta F)}) + \beta(E + F\ln((\frac{\beta F}{(1+\beta F)})k_t^{\theta}))) \tag{23}$$

Vemos el detalle que la Value Function descripta en (23) ya no lleva la palabra Max. La razón es que como estamos cargando en la Value Function las leyes optimas condicionales, esto ya es un Max. Muy bien, vamos a hacer un poco de álgebra para identificar la forma funcional implícita en la última ecuación de (23): Para ello, expandamos los logaritmos:

$$V(k_t) = \theta\ln(k_t) - \ln(1+\beta F) + \beta E + \beta F\theta\ln(k_t) + \beta F\ln(\frac{\beta F}{1+\beta F}) \tag{24}$$

Reagrupando términos, tenemos:

$$V(k_t) = (-\ln(1+\beta F) + \beta E + \beta F\ln(\frac{\beta F}{1+\beta F})) + (\theta + \theta\beta F)\ln(k_t) \tag{25}$$

Como vemos, el primer término de (25) es una constante. De esta forma, la forma funcional de la Value Function correspondiente a (25) es:

$$V(k_t) = C_1 + C_2\ln(k_t) \tag{26}$$

Recordemos que el guess de la Value Function era:

$$V(k_t) = E + F\ln(k_t) \tag{27}$$

Como vemos, (26) y (27) tienen ***EXACTAMENTE*** la ***MISMA FORMA FUNCIONAL***. Esto implica que el guess de la Value Function es ***TOTALMENTE CONSISTENTE*** con la esencia del modelo y de esta forma nos permitirá encontrar la forma reducida de las leyes óptimas del sistema.

<u>Paso 6: Determinación de la Convergencia Paramétrica de la Value Function</u>

Muy bien, si el guess de la Value Function pasó el test de autorreplicación funcional, ahora lo único que falta es hacer converger a los parámetros C_1, C_2 de la Value Function. Es aquí en donde se hace necesario el concepto de ***CONTRACTION MAPPING***. Este concepto implica que la Value Function debería iterar en sus parámetros desconocidos C_1, C_2 hasta que los mismos no cambien mas. En este punto de no cambio, la Value Function se contrae al punto que no cambia mas. Es en este punto en donde los parámetros desconocidos C_1, C_2 se identifican. Lo importante de identificar a los parámetros C_1, C_2 es que una vez conocidos los mismos podemos determinar finalmente las policy functions del sistema.

Para entender qué implica la iteración de la Value Function, analicemos la siguiente secuencia de eventos:

1) La iteración comienza con parámetros E, F;
2) Seguimos los pasos (23) a (27) y determinamos los nuevos parámetros C_1, C_2;
3) Volvemos al paso (23) en donde ahora $E = C_1$ y $F = C_2$;
4) Seguimos con los pasos (23) a (27) y determinamos los nuevos parámetros C_1', C_2';
5) El proceso continua hasta que $C_1 = C_1'$ y $C_2 = C_2'$.

En nuestro caso particular, la convergencia de los parámetros se obtiene de la siguiente forma:

$$F = \theta + \theta\beta F \Rightarrow F(1-\theta\beta) = \theta \Rightarrow F = \frac{\theta}{(1-\theta\beta)} \qquad (28)$$

Respecto a la constante E la misma involucra un álgebra complica y de todas formas no afecta en absoluto a las policy functions del sistema (recordemos que en (22) solamente aparece el parámetro F afectando a las policy functions). Por lo tanto, no calcularemos el valor convergente de E.

Volvamos ahora a (22) y concentrémonos en la policy function del consumo:

$$c_t = \frac{k_t^{\theta}}{(1+\beta F)} \qquad (29)$$

Recordemos que:

$$(1+\beta F) = 1 + \beta(\frac{\theta}{(1-\theta\beta)}) = \frac{1-\theta\beta+\theta\beta}{(1-\theta\beta)} = \frac{1}{(1-\theta\beta)} \qquad (30)$$

De esta forma, (29) queda expresado como:

$$c_t = k_t^{\theta}(1-\theta\beta) \qquad (31)$$

Es muy importante detenernos en (31). Precisamente, (31) indica la solución del sistema en el sentido que despejamos la policy function como función de parámetros exógenos mas la variable de estado. De esta forma, (31) es la solución del sistema y nos indica la secuencia óptima de consumo intertemporal.

Obviamente, podemos ahora encontrar como contrapartida obvia la ley óptima de acumulación del stock de capital:

$$k_{t+1} = (\frac{\beta F}{(1+\beta F)})k_t^{\theta} \qquad (32)$$

Recordemos que:

$$(\frac{\beta F}{(1+\beta F)}) = \beta F \frac{1}{(1+\beta F)} = \frac{\beta\theta}{(1-\theta\beta)}(1-\theta\beta) \Rightarrow (\frac{\beta F}{(1+\beta F)}) = \beta\theta \qquad (33)$$

Utilizando (33) en (32), tenemos:

$$k_{t+1} = \beta\theta k_t^{\theta} \qquad (34)$$

De esta forma, la solución del sistema es:

$$c_t = k_t^{\theta}(1-\theta\beta) \qquad\qquad k_{t+1} = \beta\theta k_t^{\theta} \qquad (35)$$

Recordemos que este es el Optimal Growth Model y que ya determinamos antes que el Optimal Growth Model exhibe un equilibrio dinámico con Saddle Path en el sentido que solamente existe una secuencia de todas las posibles que alcanza un sendero estable de equilibrio. Precisamente, las ecuaciones descriptas en (35) son las ecuaciones del Saddle Path del Modelo. Lo que (35) nos está indicando es conceptualmente muy poderoso en el sentido que nos dice que es óptimo ***"subirse"*** al saddle. Uno podría preguntarse qué condición de optimalidad se violaría si la secuencia escogida no fuera el saddle. Sin entrar en detalles, la única condición que se violaría seria la Transversality Condition.

Conclusión: La secuencia óptima del modelo es la ecuación del saddle path.

Como conclusión, es muy importante sistematizar conceptualmente lo que hemos realizado:

1) Dado que el problema es no lineal, la única forma de resolverlo es a traves de un guess de la Value Function;
2) Una vez hecho el guess, es necesario calcular las condiciones de primer orden y despejar de esta forma las Euler Equations Condicionales en el guess;
3) Si se cargan estas leyes óptimas condicionales en la Value Function, esta Value Function resultante debería autorreplicar el guess de la Value Function si el mismo es consistente con el modelo. Si esto no ocurre, el guess fue incorrecto y hay que comenzar todo nuevamente con un nuevo guess;
4) Una vez encontrado un guess consistente, es necesario encontrar los parámetros de la Value Function por convergencia iterativa;
5) La convergencia iterativa de los parámetros es necesaria para poder identificar la solución para las policy functions del sistema;
6) Una vez encontrados estos parámetros, volvemos a las ecuaciones de las policy functions y resolvemos el modelo.

Finalmente, es muy importante destacar que toda esta metodología funciona ***UNICAMENTE*** bajo el supuesto de funciones de utilidad logarítmicas.

Guess de la Value Function en Problemas No Lineales, Estocásticos y con Una Sola Variable de Control

Por primera vez introduciremos al análisis una fuente de incertidumbre. Supondremos que la economía recibirá en cada periodo un shock tecnológico. El shock le ***PEGA*** a la función de producción de la siguiente forma:

$$f(k_t) = k_t^{\theta} e^{z_t} \qquad (1)$$

Obviamente, lo que genera incertidumbre en este modelo es el término z_t. Vamos a analizar cual es la ecuación en diferencias que describe la dinámica de la variable z_t:

$$z_t = \gamma z_{t-1} + \varepsilon_t \qquad (2)$$

Es importante destacar que esta ecuación en diferencias de primer orden se diferencia de todas las que analizamos antes en el sentido que es estocástica. La fuente estocástica de la ecuación descripta en (2) está explicada por el término ε_t. Este término se distribuye de la siguiente forma:

$$\varepsilon_t \sim N(0,\sigma) \qquad (3)$$

Para que el modelo sea convergente, es necesario que $abs(\gamma) < 1$.

Paso 1: Definición del Problema de Optimización Dinámica

El problema de optimización en este caso es:

$$\begin{aligned} & Max \sum_{t=0}^{\infty} \beta^t \ln(c_t) \qquad \text{st :} \\ & k_{t+1} = k_t^{\theta} e^{z_t} - c_t \\ & z_{t+1} = \gamma z_t + \varepsilon_{t+1} \qquad \varepsilon_{t+1} \sim N(0,\sigma) \\ & k_0 \ dado \\ & Transversality\ Condition : \lim_{t\to\infty} \beta^t u'(c_t) k_t = 0 \end{aligned} \qquad (4)$$

Vemos que lo único que lo diferencia del modelo son dos aspectos. Primero, la función de producción está constantemente golpeada por un shock estocástico: $k_t^{\theta} e^{z_t}$. Segundo, el modelo tiene una ecuación en diferencias adicional que describe la dinámica de la variable estocástica z_t.

Paso 2: Cálculo del Certainty Version of the Steady State

Obviamente, un modelo que constantemente en cada periodo recibe un shock estocástico ***NUNCA*** puede alcanzar el steady state. Sin embargo, para tener una idea de adonde esta economía se ubicaría en promedio se define el ***"certainty version of the steady state"***. Lo que se hace en estos casos es extremadamente sencillo: se anula el shock estocástico haciendo que $z_t = 0 \, \forall t$. De esta forma, el problema definido en (4) es idéntico al caso anterior de no incertidumbre, por lo cual, el certainty version del steady state coincide exactamente con el caso anterior:

$$\bar{k} = \left(\frac{1}{\beta\theta}\right)^{1/(\theta-1)} \qquad \bar{c} = \left(\frac{1}{\beta\theta}\right)^{\theta/(\theta-1)} - \left(\frac{1}{\beta\theta}\right)^{1/(\theta-1)} \tag{5}$$

Antes de continuar, es muy importante determinar qué es estocástico en el periodo t. En el periodo t, tanto k_t como z_t son conocidas al inicio del periodo. Esto hace que tanto k_t como z_t sean variables de estado en el periodo t. Con lo cual, la única fuente de incertidumbre entra en el periodo próximo a t. Veamos qué significa esto con mas precisión. La ecuación de Bellman en este caso es:

$$V(k_t, z_t) = Max(\ln(c_t) + \beta E_t(V(k_{t+1}))) = Max(\ln(c_t) + \beta E_t(V(k_t^{\theta} e^{z_t} - c_t))) \tag{6}$$

En primer lugar, observamos que ahora lo que se maximiza es ***el valor esperado*** de la utilidad intertemporal futura ***condicional*** en lo que conocemos en t (de ahí el subíndice t en la esperanza matemática). Vemos por otra parte que dado que z_t es una variable de estado en t, la única fuente de incertidumbre en (6) está contenida implícitamente en el segundo término de la ecuación de Bellman: $\beta E_t(V(k_{t+1}))$. La intuición de esto es la siguiente: dado que espero un shock estocástico mañana (z_{t+1}), el consumo que determino hoy (c_t) tendrá que se optimo respecto a esa expectativa de shock. Es decir, k_{t+1} tendrá que ser tal que ***SOPORTE*** óptimamente el shock que la función de producción recibirá en $t+1$: $f(k_{t+1}) = k_{t+1}^{\theta} e^{z_{t+1}}$. Con lo cual, este problema de optimización intertemporal estocástico ***ANTICIPA*** el shock futuro a traves de su esperanza matemática.

La esperanza matemática en t del shock en $t+1$ es:

$$E_t(z_{t+1}) = E_t(\gamma z_t + \varepsilon_{t+1}) = E_t(\gamma z_t) + E_t(\varepsilon_{t+1}) = \gamma z_t \tag{7}$$

Es este término el que implícitamente está contenido en $\beta E_t(V(k_{t+1}))$ y es el que diferencia a este problema de uno de total certidumbre. Finalmente, ahora la Value Function se expresa como $V(k_t, z_t)$, es decir, depende ahora de dos variables de estado. Por otra parte, el hecho de que $V(k_t, z_t)$ sea función del término z_t nos indica que la optimalidad dinámica será función al menos en parte del término z_t que no solamente impacta a la función objetivo en t sino que se propaga en $t+1$ a traves de γz_t.

<u>Paso 3: Guess de la Value Function</u>

Dada la alta no linearidad del problema, otra vez debemos realizar un guess de la Value Function:

$$V(k_t, z_t) = E + F\ln(k_t) + Gz_t \tag{8}$$

De esta forma, ***CONDICIONAL*** en el guess, la ecuación de Bellman es:

$$V(k_t, z_t) = Max\ln(\mathrm{c_t}) + \beta E_t(V(k_{t+1})) = Max\ln(\mathrm{c_t}) + \beta E_t(E + F\ln(k_{t+1}) + Gz_{t+1}) \tag{9}$$

Recordemos que:

$$k_{t+1} = k_t^{\theta} e^{z_t} - c_t \tag{10}$$

De esta forma, la ecuación de Bellman resulta:

$$V(k_t, z_t) = Max \ln(\mathrm{c_t}) + \beta E_t (E + F \ln(k_t^{\theta} e^{z_t} - c_t) + Gz_{t+1}) \tag{11}$$

Vemos que la incertidumbre futura del modelo está contenida en el termino: Gz_{t+1}. Recordemos que z_{t+1} es desconocido en t. La pregunta que nos podemos hacer es: qué tan desconocido es z_{t+1} en t?. Bueno, la respuesta es que dado que $E_t(z_{t+1}) = \gamma z_t$, esto nos dice que el ***PRESENTE*** (representado por el término z_t) nos está informando bastante sobre el ***FUTURO***. Es en este sentido en donde podemos decir que z_{t+1} no es totalmente desconocido en t. Otra pregunta que nos podemos hacer es: por qué esto es así?. La respuesta reside en la relación dinámica que supusimos entre z_t y z_{t+1}. Ambas variables se suponemos correlacionadas en el tiempo de acuerdo a un AR(1) process. Es por esta única razón que el estado en t de la variable z me permite inferir algo respecto al estado de esta variable en $t+1$.

De esta forma, el término $E_t(E + F \ln(k_{t+1}) + Gz_{t+1})$ es igual a:

$$E_t(E + F \ln(k_t^{\theta} e^{z_t} - c_t) + Gz_{t+1}) = E + F \ln(k_t^{\theta} e^{z_t} - c_t) + G\gamma z_t \tag{12}$$

Con lo cual, la Bellman Equation es:

$$V(k_t, z_t) = Max \ln(\mathrm{c_t}) + E + F \ln(k_t^{\theta} e^{z_t} - c_t) + G\gamma z_t \tag{13}$$

<u>Paso 4: Cálculo de la Condición de Primer Orden y Leyes Optimas Condicionales</u>

$$\frac{\partial V(k_t, z_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1}, z_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \qquad \frac{\partial k_{t+1}}{\partial c_t} = -1 \tag{14}$$

Ahora, dado el guess de la Value Function, contamos con un dato adicional que antes desconocíamos. Este dato adicional está representado por el hecho que ahora sabemos qué es $\partial V(k_{t+1}, z_{t+1}) / \partial k_{t+1}$:

$$\frac{\partial V(k_{t+1}, z_{t+1})}{\partial k_{t+1}} = \frac{\partial (E + F \ln(k_t^{\theta} e^{z_t} - c_t))}{\partial k_{t+1}} = \frac{F}{(k_t^{\theta} e^{z_t} - c_t)} \tag{15}$$

Poniendo (15) en (14) tenemos:

$$\frac{\partial V(k_t, z_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{F}{(k_t^{\theta} e^{z_t} - c_t)}(-1) = 0 \tag{16}$$

Esto a su vez implica:

$$\frac{1}{c_t} = \beta \frac{F}{(k_t^{\theta} e^{z_t} - c_t)} \Rightarrow (k_t^{\theta} e^{z_t} - c_t) = c_t \beta F \Rightarrow c_t(1 + \beta F) = k_t^{\theta} e^{z_t} \Rightarrow \tag{17}$$

$$c_t = \frac{k_t^{\theta} e^{z_t}}{(1 + \beta F)}$$

Es muy importante detenernos en la expresión (17). Vemos que esta expresión es la policy function del consumo ***CONDICIONAL*** en el guess de la Value Function. Si el guess de la Value Function fuera erróneo, esta ley óptima no tendría nada que ver con la verdadera ley óptima del sistema. De la misma forma, la ley óptima para la acumulación de capital es:

$$k_{t+1} = k_t^{\theta} e^{z_t} - c_t \Rightarrow k_{t+1} = k_t^{\theta} e^{z_t} - \frac{k_t^{\theta} e^{z_t}}{(1+\beta F)} \Rightarrow \tag{18}$$

$$k_{t+1} = \frac{k_t^{\theta} e^{z_t} + \beta F k_t^{\theta} e^{z_t} - k_t^{\theta} e^{z_t}}{(1+\beta F)} \Rightarrow k_{t+1} = (\frac{\beta F}{(1+\beta F)}) k_t^{\theta} e^{z_t}$$

Muy bien, las dos leyes optimas ***CONDICIONALES*** del sistema son:

$$c_t = \frac{k_t^{\theta} e^{z_t}}{(1+\beta F)} \qquad k_{t+1} = (\frac{\beta F}{(1+\beta F)}) k_t^{\theta} e^{z_t} \tag{19}$$

<u>Paso 5: Determinación de la Replicación Funcional de la Value Function</u>

Si el guess de la Value Function fue correcto, si cargásemos las leyes optimas condicionales descriptas en (19) en la Value Function supuesta, la misma debería ***AUTORREPLICARSE***. Vamos a determinar si esto en verdad ocurre:

$$V(k_t, z_t) = \ln(\frac{k_t^{\theta} e^{z_t}}{(1+\beta F)}) + \beta (E + F \ln((\frac{\beta F}{(1+\beta F)}) k_t^{\theta} e^{z_t})) + G\gamma z_t) \tag{20}$$

Vemos el detalle que la Value Function descripta en (20) ya no lleva la palabra Max. La razón es que como estamos cargando en la Value Function las leyes optimas condicionales, esto ya es un Max. Muy bien, vamos a hacer un poco de álgebra para identificar la forma funcional implícita en la última ecuación de (20). Para ello, expandamos los logaritmos:

$$V(k_t, z_t) = \theta \ln(k_t) + z_t - \ln(1+\beta F) + \beta E + \beta F \theta \ln(k_t) + \beta F z_t + \beta F \ln(\frac{\beta F}{1+\beta F}) + \beta G \gamma z_t \tag{21}$$

Reagrupando términos, tenemos:

$$V(k_t, z_t) = (-\ln(1+\beta F) + \beta E + \beta F \ln(\frac{\beta F}{1+\beta F})) + (\theta + \theta\beta F)\ln(k_t) + (1+\beta F + \beta G \gamma) z_t \tag{22}$$

Como vemos, el primer término de (22) es una constante. De esta forma, la forma funcional de la Value Function correspondiente a (22) es:

$$V(k_t, z_t) = C_1 + C_2 \ln(k_t) + C_3 z_t \tag{23}$$

Recordemos que el guess de la Value Function era:

$$V(k_t, z_t) = E + F \ln(k_t) + G z_t \tag{24}$$

Como vemos, (23) y (24) tienen ***EXACTAMENTE*** la ***MISMA FORMA FUNCIONAL***. Esto implica que el guess de la Value Function es ***TOTALMENTE CONSISTENTE*** con la esencia del modelo y de esta forma nos permitirá encontrar la forma reducida de las leyes óptimas del sistema.

<u>Paso 6: Determinación de la Convergencia Paramétrica de la Value Function</u>

Muy bien, si el guess de la Value Function pasó el test de autorreplicación funcional, ahora lo único que falta es hacer converger a los parámetros C_1, C_2, C_3 de la Value Function. Es aquí en donde se hace necesario el concepto de ***CONTRACTION MAPPING***. En nuestro caso particular, la convergencia de los parámetros se obtiene de la siguiente forma:

$$F = \theta + \theta\beta F \Rightarrow F(1-\theta\beta) = \theta \Rightarrow F = \frac{\theta}{(1-\theta\beta)} \tag{25}$$

$$G = (1 + \beta F + \beta G\gamma) \Rightarrow G(1 + \beta\gamma) = (1 + \beta F) \Rightarrow G = \frac{(1 + \beta F)}{(1 + \beta\gamma)} \tag{26}$$

Relacionando (26) y (25), tenemos:

$$G = (1 + \beta F + \beta G\gamma) \Rightarrow G(1 + \beta\gamma) = (1 + \beta F) \Rightarrow G = \frac{(1 + \frac{\beta\theta}{(1-\theta\beta)})}{(1 + \beta\gamma)} \tag{27}$$

Simplificando:

$$G = \frac{(\frac{1-\theta\beta+\theta\beta}{(1-\theta\beta)})}{(1+\beta\gamma)} = \frac{1}{(1-\theta\beta)(1+\beta\gamma)} \tag{28}$$

Igual que en el modelo anterior, respecto a la constante E la misma involucra un álgebra complica y de todas formas no afecta en absoluto a las policy functions del sistema (recordemos que en (19) solamente aparecen los parámetros F, G afectando a las policy functions). Por lo tanto, no calcularemos el valor convergente de E.

Volvamos ahora a (19) y concentrémonos en la policy function del consumo:

$$c_t = \frac{k_t^\theta e^{z_t}}{(1+\beta F)} \tag{29}$$

Recordemos que:

$$(1+\beta F) = 1 + \beta(\frac{\theta}{(1-\theta\beta)}) = \frac{1-\theta\beta+\theta\beta}{(1-\theta\beta)} = \frac{1}{(1-\theta\beta)} \tag{30}$$

De esta forma, (29) queda expresado como:

$$c_t = k_t^\theta e^{z_t}(1-\theta\beta) \tag{31}$$

Es muy importante detenernos en (31). Precisamente, (31) indica la solución del sistema en el sentido que despejamos la policy function como función de parámetros exógenos mas la variable de estado. De esta forma, (31) es la solución del sistema y nos indica la secuencia óptima de consumo intertemporal.

Obviamente, podemos ahora encontrar como contrapartida obvia la ley óptima de acumulación del stock de capital:

$$k_{t+1} = (\frac{\beta F}{(1+\beta F)})k_t^\theta e^{z_t} \tag{32}$$

Recordemos que:

$$(\frac{\beta F}{(1+\beta F)}) = \beta F \frac{1}{(1+\beta F)} = \frac{\beta\theta}{(1-\theta\beta)}(1-\theta\beta) \Rightarrow (\frac{\beta F}{(1+\beta F)}) = \beta\theta \tag{33}$$

Utilizando (33) en (32), tenemos:

$$k_{t+1} = \beta\theta k_t^\theta e^{z_t} \tag{34}$$

De esta forma, la solución del sistema estocástico es:

$$c_t = k_t^{\theta} e^{z_t}(1-\theta\beta) \qquad\qquad k_{t+1} = \beta\theta k_t^{\theta} e^{z_t} \tag{35}$$

Como vemos, las leyes óptimas solamente quedan como función de las variables de estado en $t: k_t, z_t$, lo cual implica que solucionamos el problema de optimización intertemporal estocástico. Otra vez, nos podemos preguntar en dónde está implícita la incertidumbre futura. En realidad, las leyes óptimas nos dicen que el consumo en t anticipa al shock futuro z_{t+1} a traves del valor en t de la variable z. De esta forma, z_t se convierte en el predictor de z_{t+1} y es la forma en la que la incertidumbre futura afecta la optimalidad en t.

De esta forma, si en el periodo t la economía recibió un shock muy favorable, la policy function del consumo me indica que consuma mucho ya que no será necesario ***"preparar"*** al stock de capital k_{t+1} para un shock negativo en $t+1$, dado que el valor de z_t nos estaría indicando que la probabilidad de que eso ocurra es muy baja (recordemos que z_{t+1} está correlacionado con z_t a traves del parámetro γ).

De esta forma, toda la incertidumbre futura está contenida exclusivamente en el término e^{z_t}. Si bien la variable z_t es conocida con total certidumbre en t lo cual la convierte en una variable de estado, la misma puede verse además como el único predictor de la incertidumbre que se viene. Con lo cual, la única forma en la que la incertidumbre futura está presente en la policy function del consumo es a traves del término e^{z_t}. De esta forma, z_t anticipa a traves de su valor, el valor de z_{t+1}. Lo que la policy function del consumo nos dice respecto a cómo consumir óptimamente es que si el shock en t fue muy positivo, es óptimo consumir mucho por dos razones: 1) porque tengo mucho para consumir, 2) porque para el periodo siguiente no tengo que generar un gran stock de capital que me proteja contra un shock muy negativo, dado que espero que el shock sea muy parecido al actual.

Finalmente es importante diferenciar las policy functions que resultan del caso con total certidumbre respecto al caso con certidumbre. En total certidumbre, una vez determinada la secuencia óptima, la misma no cambia nunca mas dado que el modelo no ofrece absolutamente ningún cambio. Por el contrario, en el caso con incertidumbre, en cada periodo de tiempo hay que **AJUSTAR** la policy function de acuerdo a la nueva realización de la variable estocástica z. Es decir, en este caso la policy function se construye a medida que transcurre el tiempo y de acuerdo a la magnitud del shock contenido en z_t. Es decir, es imposible determinar en t toda la secuencia hasta el infinito.

Guess de la Value Function en Problemas No Lineales, Estocásticos y con Dos Variables de Control

Paso 1: Definición del Problema de Optimización Dinámica

El problema de optimización en este caso es:

$$Max \sum_{t=0}^{\infty} \beta^t (\ln(c_t) - Ah_t) \qquad \text{st}: \tag{1}$$

$$k_{t+1} = k_t^{\theta} h_t^{(1-\theta)} e^{z_t} - c_t$$

$$z_{t+1} = \gamma z_t + \varepsilon_{t+1} \qquad \varepsilon_{t+1} \sim N(0,\sigma)$$

$$k_0 \; dado$$

$$Transversality\ Condition: \lim_{t\to\infty} \beta^t u^{'}(c_t) k_t = 0$$

$$Transversality\ Condition: \lim_{t\to\infty} \beta^t u^{'}(h_t) k_t = 0$$

Vemos que este modelo exhibe dos diferencias respecto al anterior. Primero, las horas trabajadas entran en la función de utilidad, generando desutilidad. Segundo, las horas trabajadas entran en la función de producción. La ecuación de Bellman es:

$$V(k_t, z_t) = Max_{c_t, h_t} \ \ln(c_t - Ah_t) + \beta E_t (V(k_t^{\theta} h_t^{(1-\theta)} e^{z_t} - c_t)) \qquad (2)$$

Paso 2: Cálculo del Certainty Version of the Steady State

En la certainty equivalent version del steady state, anulamos el termino estocástico y por lo tanto la Bellman Equation resulta:

$$V(k_t) = Max_{c_t, h_t} \ \ln(c_t - Ah_t) + \beta (V(k_t^{\theta} h_t^{(1-\theta)} e^{z_t} - c_t)) \qquad (3)$$

Dado que hay dos variables de control, es necesario calcular dos condiciones de primer orden:

$$\frac{\partial V(k_t)}{\partial c_t} = 0 \qquad \frac{\partial V(k_t)}{\partial h_t} = 0 \qquad (4)$$

Comencemos con la primera:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \quad \frac{\partial k_{t+1}}{\partial c_t} = -1 \qquad (5)$$

La condición de primer orden respecto a la variable de control es:

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial c_t} = 0 \quad \frac{\partial k_{t+1}}{\partial c_t} = -1 \qquad (6)$$

$$\frac{\partial V(k_t)}{\partial c_t} = \frac{1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} (-1) = 0 \qquad (7)$$

Esta condición de primer orden implica:

$$\frac{1}{c_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \Rightarrow \frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{1}{c_t \beta} \qquad (8)$$

$$\frac{1}{c_{t-1}} = \beta \frac{\partial V(k_t)}{\partial k_t} \Rightarrow \frac{\partial V(k_t)}{\partial k_t} = \frac{1}{c_{t-1} \beta}$$

Ahora, analicemos la condición de primer orden respecto a la variable de control, trabajo:

$$\frac{\partial V(k_t)}{\partial h_t} = -A + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial h_t} = 0 \quad \frac{\partial k_{t+1}}{\partial h_t} = (1-\theta) h_t^{-\theta} k_t^{\theta} \qquad (9)$$

$$\frac{\partial V(k_t)}{\partial h_t} = -A + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} (1-\theta) h_t^{-\theta} k_t^{\theta} = 0 \qquad (10)$$

$$\beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} (1-\theta) (\frac{k_t}{h_t})^{\theta} = A \Rightarrow \frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{A}{(1-\theta)\beta} (\frac{h_t}{k_t})^{\theta} \qquad (11)$$

Esta condición de primer orden implica:

$$\frac{\partial V(k_{t+1})}{\partial k_{t+1}} = \frac{A}{(1-\theta)\beta}(\frac{h_t}{k_t})^{\theta} \qquad (12)$$

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{A}{(1-\theta)\beta}(\frac{h_{t-1}}{k_{t-1}})^{\theta}$$

Derivemos ahora respecto a la variable de estado, aplicando ***Benveniste and Sheikman***:

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t} \qquad \frac{\partial k_{t+1}}{\partial k_t} = \theta k_t^{(\theta-1)} h_t^{(1-\theta)} \Rightarrow \qquad (13)$$

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \theta k_t^{(\theta-1)} h_t^{(1-\theta)}$$

Combinemos (13) y (11):

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \theta k_t^{(\theta-1)} h_t^{(1-\theta)} = \beta(\frac{A}{(1-\theta)\beta}(\frac{h_t}{k_t})^{\theta})\theta k_t^{(\theta-1)} h_t^{(1-\theta)} \qquad (14)$$

Simplificando, tenemos:

$$\frac{\partial V(k_t)}{\partial k_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} \theta k_t^{(\theta-1)} h_t^{(1-\theta)} \Rightarrow \qquad (15)$$

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{A\theta}{(1-\theta)} \frac{h_t}{k_t}$$

Es muy importante detenernos en (15). Destaquemos que hasta el momento no utilizamos ninguna de las dos condiciones de primer orden: $\partial V(k_t)/\partial c_t$, $\partial V(k_t)/\partial h_t$.

Primero, recordemos que $\partial V(k_t)/\partial c_t = 0$ implica la siguiente relación (ver (8)):

$$\frac{\partial V(k_t)}{\partial c_t} = 0 \Rightarrow \qquad (16)$$

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{1}{c_{t-1}\beta}$$

Segundo, recordemos que $\partial V(k_t)/\partial h_t = 0$ implica la siguiente relación (ver (12)):

$$\frac{\partial V(k_t)}{\partial h_t} = 0 \Rightarrow \qquad (17)$$

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{A}{(1-\theta)\beta}(\frac{h_{t-1}}{k_{t-1}})^{\theta}$$

Como vemos, (16) y (17) son los valores que surgen de las dos condiciones de optimalidad del problema y todavía no las utilizamos. Es decir, (16) y (17) tienen implícita dos Euler Equations. Llegó el momento de utilizarlas.

Primero, combinemos (15) y (16):

$$\frac{1}{c_{t-1}\beta} = \frac{A\theta}{(1-\theta)} \frac{h_t}{k_t} \qquad (18)$$

Segundo, combinemos (15) y (17):

$$\frac{A}{(1-\theta)\beta}(\frac{h_{t-1}}{k_{t-1}})^{\theta} = \frac{A\theta}{(1-\theta)}\frac{h_t}{k_t} \tag{19}$$

Finalmente, recordemos que tenemos la law of motion del capital (evaluada en la certainty version):

$$k_{t+1} = k_t^{\theta} h_t^{(1-\theta)} - c_t \tag{20}$$

Con lo cual, estas son las tres (en vez de dos) leyes dinámicas del sistema, que junto con las dos condiciones de transversalidad dominaran la dinámica del modelo. Destaquemos que (18) y (19) son las dos Euler Equations del sistema.

Recordemos que nuestro objetivo es calcular la certainty version del steady state. Con lo cual, si resolvemos el sistema de ecuación descripto por (18), (19), y (20) evaluados en steady state, tenemos:

$$\frac{1}{\overline{c}\beta} = \frac{A\theta}{(1-\theta)}\frac{\overline{h}}{\overline{k}} \tag{21}$$

$$\frac{A}{(1-\theta)\beta}(\frac{\overline{h}}{\overline{k}})^{\theta} = \frac{A\theta}{(1-\theta)}\frac{\overline{h}}{\overline{k}} \tag{22}$$

$$\overline{k} = \overline{k}^{\theta}\overline{h}^{(1-\theta)} - \overline{c} \tag{23}$$

Utilicemos (23) para despejar $\overline{c}$ en función de $\overline{k}$:

$$\overline{c} = \overline{k}^{\theta}\overline{h}^{(1-\theta)} - \overline{k} \tag{24}$$

Utilicemos (22) para despejar $\overline{h}$ en función de $\overline{k}$:

$$\frac{1}{\theta\beta}(\frac{\overline{h}}{\overline{k}})^{(\theta-1)} = 1 \Rightarrow \overline{h}^{(\theta-1)} = \theta\beta\overline{k}^{(\theta-1)} \Rightarrow \tag{25}$$

$$\overline{h} = \theta\beta^{(1/(\theta-1))}\overline{k}$$

Ahora combinemos (25) en (24) para despejar a $\overline{k}$ en función de $\overline{c}$:

$$\overline{c} = \overline{k}^{\theta}(\theta\beta^{(1/(\theta-1))}\overline{k})^{(1-\theta)} - \overline{k} \Rightarrow \tag{26}$$

$$\overline{c} = \overline{k}(\theta\beta)^{-1} - \overline{k} \Rightarrow \overline{c} = \overline{k}(\frac{1}{\theta\beta}) - \overline{k} \Rightarrow \overline{c} = \frac{(1-\theta\beta)}{\theta\beta}\overline{k} \Rightarrow$$

$$\overline{k} = \frac{\theta\beta}{(1-\theta\beta)}\overline{c}$$

Ahora combinemos (26) en (25) para despejar a $\overline{h}$ en función de $\overline{c}$ recordando que:

$$\frac{1}{(\theta-1)} + 1 = \frac{1+\theta-1}{(\theta-1)} = \frac{\theta}{(\theta-1)} \tag{27}$$

$$\overline{h} = \theta\beta^{(1/(\theta-1))}\overline{k} \Rightarrow \overline{h} = \theta\beta^{(1/(\theta-1))}(\overline{c}\frac{\theta\beta}{(1-\theta\beta)}) \Rightarrow \quad (28)$$

$$\overline{h} = \frac{\theta\beta^{\frac{\theta}{(\theta-1)}}}{(1-\theta\beta)}\overline{c}$$

Recordemos que:

$$\frac{\theta}{(\theta-1)} - 1 = \frac{\theta-\theta+1}{(\theta-1)} = \frac{1}{(\theta-1)} \quad (29)$$

Carguemos (26) y (27) en (21):

$$\frac{1}{\overline{c}\beta} = \frac{A\theta}{(1-\theta)}\frac{\overline{h}}{\overline{k}} \Rightarrow \frac{1}{\overline{c}\beta} = \frac{A\theta}{(1-\theta)}(\frac{\theta\beta^{\frac{\theta}{(\theta-1)}}}{(1-\theta\beta)}\overline{c})(\frac{(1-\theta\beta)}{\theta\beta\overline{c}}) \Rightarrow \quad (30)$$

$$\frac{1}{\overline{c}\beta} = \frac{A\theta}{(1-\theta)}\theta\beta^{\frac{1}{(\theta-1)}} \Rightarrow \overline{c} = \frac{1}{A}\frac{(1-\theta)}{\theta\beta}\theta\beta^{\frac{1}{(1-\theta)}} \Rightarrow$$

$$\overline{c} = \frac{(1-\theta)}{A}\theta\beta^{\frac{\theta}{(1-\theta)}} \qquad \textit{Steady State Consumption}$$

Utilizando (30) en (26) y (28), obtenemos los steady states del capital y horas trabajadas respectivamente:

$$\overline{k} = \frac{(1-\theta)\theta\beta^{1/(1-\theta)}}{A(1-\theta\beta)} \qquad \overline{h} = \frac{(1-\theta)}{A(1-\theta\beta)} \quad (31)$$

Paso 3: Guess de la Value Function

Como este también es un problema no lineal, debemos hacer un guess de la Value Function. Utilicemos el mismo guess del modelo anterior:

$$V(k_t, z_t) = E + F\ln(k_t) + Gz_t \quad (32)$$

La ecuación de Bellman es:

$$V(k_t, z_t) = Max_{c_t, h_t}\ \ln(c_t) - Ah_t + \beta E_t(E + F\ln(k_t^{\theta}e^{z_t} - c_t) + Gz_{t+1}) \quad (33)$$

Recordemos que:

$$E_t(E + F\ln(k_t^{\theta}e^{z_t}) + Gz_{t+1}) = E + F\ln(k_t^{\theta}e^{z_t} - c_t) + G\gamma z_t \quad (34)$$

Con lo cual, la Bellman Equation resulta:

$$V(k_t, z_t) = Max_{c_t, h_t}\ \ln(c_t) - Ah_t + \beta(E + F\ln(k_t^{\theta}e^{z_t} - c_t) + G\gamma z_t) \quad (35)$$

Paso 4: Cálculo de la Condición de Primer Orden y Leyes Optimas Condicionales

Consideremos la condición de primer orden respecto a la variable de control consumo:

$$\frac{\partial V(k_t, z_t)}{\partial c_t} = \frac{1}{c_t} + \beta\frac{\partial V(k_{t+1}, z_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial c_t} = 0 \quad \frac{\partial k_{t+1}}{\partial c_t} = -1 \quad (36)$$

Ahora, dado el guess de la Value Function, contamos con un dato adicional que antes desconocíamos. Este dato adicional está representado por el hecho que ahora sabemos qué es $\partial V(k_{t+1}, z_{t+1}) / \partial k_{t+1}$:

$$\frac{\partial V(k_{t+1}, z_{t+1})}{\partial k_{t+1}} = \frac{\partial (E + F \ln(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t))}{\partial k_{t+1}} = \frac{F}{(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t)} \tag{37}$$

Poniendo (37) en (36) tenemos:

$$\frac{\partial V(k_t, z_t)}{\partial c_t} = \frac{1}{c_t} + \frac{\beta F}{(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t)}(-1) = 0 \tag{38}$$

Esto a su vez implica:

$$\frac{1}{c_t} = \beta \frac{F}{(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t)} \Rightarrow (k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t) = c_t \beta F \Rightarrow c_t (1 + \beta F) = k_t^\theta h_t^{(1-\theta)} e^{z_t} \Rightarrow \tag{39}$$

$$c_t = \frac{k_t^\theta h_t^{(1-\theta)} e^{z_t}}{(1 + \beta F)}$$

Es muy importante detenernos en la expresión (39). Vemos que esta expresión es la policy function del consumo ***CONDICIONAL*** en el guess de la Value Function. Si el guess de la Value Function fuera erróneo, esta ley óptima no tendría nada que ver con la verdadera ley óptima del sistema.

Consideremos la condición de primer orden respecto a la variable de control horas trabajadas:

$$\frac{\partial V(k_t, z_t)}{\partial h_t} = -A + \beta \frac{\partial V(k_{t+1}, z_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial h_t} = 0 \quad \frac{\partial k_{t+1}}{\partial h_t} = (1-\theta) h_t^{-\theta} k_t^\theta e^{z_t} \tag{40}$$

$$\frac{\partial V(k_t, z_t)}{\partial h_t} = -A + \frac{\beta F}{(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t)}(1-\theta) h_t^{-\theta} k_t^\theta e^{z_t} = 0 \Rightarrow \tag{41}$$

$$A(k_t^\theta h_t^{(1-\theta)} e^{z_t} - c_t) = \beta F (1-\theta) h_t^{-\theta} k_t^\theta e^{z_t}$$

Poniendo (39) en (41), tenemos:

$$A(k_t^\theta h_t^{(1-\theta)} e^{z_t} - \frac{k_t^\theta h_t^{(1-\theta)} e^{z_t}}{(1 + \beta F)}) = \beta F (1-\theta) h_t^{-\theta} k_t^\theta e^{z_t} \tag{42}$$

Simplificando:

$$A \beta F \frac{k_t^\theta h_t^{(1-\theta)} e^{z_t}}{(1 + \beta F)} = \beta F (1-\theta) h_t^{-\theta} k_t^\theta e^{z_t} \tag{43}$$

$$\frac{\beta F(1-\theta)(1+\beta F)}{A\beta F}=\frac{h_t^{(1-\theta)}k_t^{\theta}e^{z_t}}{h_t^{-\theta}k_t^{\theta}e^{z_t}}\Rightarrow h_t=\frac{(1-\theta)(1+\beta F)}{A}\Rightarrow$$
$$h_t=T_2 \qquad (44)$$

Vemos que (44) describe la policy function de horas trabajadas condicional en el guess de la Value Function y genera una secuencia de oferta de trabajo constante en el tiempo. Ahora pongamos (44) en (39) para despejar la policy function condicional del consumo:

$$c_t=\frac{k_t^{\theta}h_t^{(1-\theta)}e^{z_t}}{(1+\beta F)}=\frac{k_t^{\theta}e^{z_t}}{(1+\beta F)}h_t^{(1-\theta)}=\frac{k_t^{\theta}e^{z_t}}{(1+\beta F)}\left(\frac{(1-\theta)(1+\beta F)}{A}\right)^{(1-\theta)}\Rightarrow$$
$$c_t=k_t^{\theta}e^{z_t}T_1 \qquad (45)$$

Vemos que en definitiva, T_1, T_2 son dos constantes. Veamos ahora si la Value Function se autorreplica:

$$V(k_t,z_t)=\ln(k_t^{\theta}e^{z_t}T_1)-AT_2+\beta(E+F\ln(k_t^{\theta}T_2^{(1-\theta)}e^{z_t}-k_t^{\theta}e^{z_t}T_1)+G\gamma z_t)\Rightarrow \qquad (46)$$
$$V(k_t,z_t)=\ln(k_t^{\theta}e^{z_t}T_1)-AT_2+\beta(E+F\ln(k_t^{\theta}e^{z_t}(T_2^{(1-\theta)}-T_1))+G\gamma z_t)$$

Ampliemos la expresión calculando logaritmos:

$$V(k_t,z_t)=\ln(T_1)+z_t+\theta\ln(k_t)-AT_2+\beta E+\beta F\ln(T_2^{(1-\theta)}-T_1)+\beta Fz_t+\theta\beta F\ln(k_t)+\beta G\gamma z_t \qquad (47)$$

Agrupando, tenemos:

$$V(k_t,z_t)=C_1+(\theta+\theta\beta F)\ln(k_t)+(1+\beta F+\beta G\gamma)z_t \qquad (48)$$

Vemos que en definitiva, la forma funcional autorreplicada es:

$$V(k_t,z_t)=C_1+C_2\ln(k_t)+C_3z_t \qquad (49)$$

Vemos que la Value Function autogenerada replica la forma funcional del guess, lo cual implica que el guess fue correcto. Ahora es necesario evaluar los parámetros de la Value Function por convergencia iterativa:

$$F=C_2\Rightarrow F=(\theta+\theta\beta F)\Rightarrow F=\frac{\theta}{(1-\theta\beta)} \qquad (50)$$

El segundo parámetro es:

$$G=C_3\Rightarrow G=(1+\beta F+\beta G\gamma)\Rightarrow G(1+\beta\gamma)=(1+\beta F)\Rightarrow G=\frac{(1+\beta F)}{(1+\beta\gamma)} \qquad (51)$$

Relacionando (51) y (50), tenemos:

$$G=(1+\beta F+\beta G\gamma)\Rightarrow G(1+\beta\gamma)=(1+\beta F)\Rightarrow G=\frac{(1+\frac{\beta\theta}{(1-\theta\beta)})}{(1+\beta\gamma)} \qquad (52)$$

Simplificando:

$$G=\frac{\left(\frac{1-\theta\beta+\theta\beta}{(1-\theta\beta)}\right)}{(1+\beta\gamma)}=\frac{1}{(1-\theta\beta)(1+\beta\gamma)} \tag{53}$$

Recordemos que:

$$1+\beta F=1+\frac{\beta\theta}{(1-\theta\beta)}=\frac{1}{(1-\theta\beta)} \tag{54}$$

Utilizando (44), la policy function de las horas trabajadas es:

$$h_t=\frac{(1-\theta)(1+\beta F)}{A}\Rightarrow h_t=\frac{(1-\theta)}{A(1-\theta\beta)} \tag{55}$$

Utilizando (45), la policy function del consumo es:

$$c_t=\frac{k_t^{\theta}e^{z_t}}{(1+\beta F)}\left(\frac{(1-\theta)(1+\beta F)}{A}\right)^{(1-\theta)}\Rightarrow c_t=k_t^{\theta}e^{z_t}(1-\theta\beta)\left(\frac{(1-\theta)}{A(1-\theta\beta)}\right)^{(1-\theta)}\Rightarrow \tag{56}$$

$$c_t=k_t^{\theta}e^{z_t}(1-\theta\beta)^{\theta}\left(\frac{(1-\theta)}{A}\right)^{(1-\theta)}$$

De esta forma, la solución del sistema es:

$$c_t=k_t^{\theta}e^{z_t}(1-\theta\beta)^{\theta}\left(\frac{(1-\theta)}{A}\right)^{(1-\theta)} \qquad\qquad h_t=\frac{(1-\theta)}{A(1-\theta\beta)} \tag{57}$$

Vemos que (57) describe las ecuaciones del saddle path que describen las secuencias intertemporales óptimas del modelo.

Sistemas No Lineales: Evaluación Algebraica de su Estabilidad Dinámica

Las técnicas que describimos anteriormente, utilizando esencialmente álgebra matricial, solo son aplicables para el caso de sistemas de ecuaciones en diferencias **LINEALES**. Para el caso de sistemas no lineales, como por ejemplo, el optimal growth model, las ecuaciones en diferencias no están expresadas linealmente respecto a parámetros. Sino que su relación es mas compleja. Recordemos el sistema de ecuaciones para este caso:

$$k_{t+1}=f(k_t)+(1-\delta)k_t-c_t \qquad\qquad c_{t+1}=c_t\beta[f'(k_{t+1})+(1-\delta)] \tag{1}$$

Vemos aquí, que es imposible identificar una matriz A y utilizar las reglas analizadas anteriormente. Sin embrago, existe una forma de hacer que este sistema quede expresado como si fuese uno lineal. Para ello es necesario linearizar el sistema utilizando una expansión de Taylor. Recordemos que para el caso de dos variables, una expansión de Taylor resulta en lo siguiente:

$$x_{t+1}=f(x_t,y_t)\Rightarrow x_{t+1}=f(\bar{x},\bar{y})+f_x(\bar{x},\bar{y})(x_t-\bar{x})+f_y(\bar{x},\bar{y})(y_t-\bar{y}) \tag{2}$$

$$y_{t+1}=g(x_t,y_t)\Rightarrow y_{t+1}=g(\bar{x},\bar{y})+g_x(\bar{x},\bar{y})(x_t-\bar{x})+g_y(\bar{x},\bar{y})(y_t-\bar{y})$$

Vemos que la expansión la realizamos en torno al steady state del sistema. Si reagrupamos términos tenemos que:

$$x_{t+1}=f(\bar{x},\bar{y})-f_x(\bar{x},\bar{y})(\bar{x})-f_y(\bar{x},\bar{y})(\bar{y})+f_x(\bar{x},\bar{y})(x_t)+f_y(\bar{x},\bar{y})(y_t) \tag{3}$$

$$y_{t+1}=g(\bar{x},\bar{y})-g_x(\bar{x},\bar{y})(\bar{x})-g_y(\bar{x},\bar{y})(\bar{y})+g_x(\bar{x},\bar{y})(x_t)+g_y(\bar{x},\bar{y})(y_t)$$

Vemos que los primeros tres términos de cada ecuación son constantes, entonces podemos reexpresar el sistema de la siguiente forma:

$$x_{t+1} = b_{11} + f_x(\bar{x},\bar{y})(x_t) + f_y(\bar{x},\bar{y})(y_t) \qquad (4)$$

$$y_{t+1} = b_{12} + g_x(\bar{x},\bar{y})(x_t) + g_y(\bar{x},\bar{y})(y_t)$$

Además, las derivadas evaluadas en el steady state no son mas que un numero, esto implica que el sistema se puede reexpresar así:

$$x_{t+1} = b_{11} + a_{11}(x_t) + a_{12}(y_t) \qquad (5)$$

$$y_{t+1} = b_{12} + a_{12}(x_t) + a_{22}(y_t)$$

Vemos que este sistema linearizado tiene exactamente la misma forma que los sistemas autónomos analizados anteriormente. Vemos además, tal como antes, que el comportamiento dinámico del sistema lo determina la matriz A (que en este caso surge de linearizar a la verdadera matriz de coeficientes). Esta matriz A surge de calcular las derivadas parciales primeras del sistema con respecto a sus variables relevantes y evaluarlas en el steady state. Una vez hecho esto podemos determinar si el sistema es estable o no.

Vamos a analizar esto en el contexto del optimal growth model suponiendo que la función de producción tiene la siguiente forma funcional:

$$f(k) = k^{\alpha} \qquad (6)$$

Entonces, la matriz de derivadas primeras (o el Jacobean) evaluadas en el steady state son:

$$\frac{\partial c_{t+1}}{\partial c_t} = \beta[f^{'}(k) + (1-\delta)] + c\beta f^{''}(k)(-1) \qquad \frac{\partial k_{t+1}}{\partial c_t} = -1 \qquad (7)$$

$$\frac{\partial c_{t+1}}{\partial k_t} = c_t \beta f''(k)[f^{'}(k) + (1-\delta)] \qquad \frac{\partial k_{t+1}}{\partial k_t} = \alpha k^{(\alpha-1)} + (1-\delta)$$

Vemos que las derivadas del consumo se hacen especialmente complicadas de calcular dado que el consumo afecta a k_{t+1}:

$$c_{t+1} = c_t \beta[f' + (1-\delta)] \Rightarrow \frac{\partial c_{t+1}}{\partial c_t} = \beta[f' + (1-\delta)] + c_t \beta[f'' \frac{\partial k_{t+1}}{\partial c_t}] \Rightarrow \qquad (8)$$

$$\frac{\partial k_{t+1}}{\partial c_t} = -1 \Rightarrow \frac{\partial c_{t+1}}{\partial c_t} = \beta[f' + (1-\delta)] + c_t \beta[f''(-1)]$$

$$c_{t+1} = c_t \beta[f' + (1-\delta)] \Rightarrow \frac{\partial c_{t+1}}{\partial k_t} = c\beta f^{''}[\frac{\partial k_{t+1}}{\partial k_t}] \Rightarrow \qquad (9)$$

$$\frac{\partial k_{t+1}}{\partial k_t} = f^{'} + (1-\delta) \Rightarrow \frac{\partial c_{t+1}}{\partial k_t} = c\beta f^{''}[f^{'} + (1-\delta)]$$

Finalmente, es necesario calcular el steady state para el capital:

$$c_{t+1} = c_t \beta[f'(k_{t+1}) + (1-\delta) \Rightarrow 1 = \beta[\alpha k_{t+1}^{\alpha-1}) + (1-\delta) \Rightarrow k = \ln[((1/\beta - (1-\delta))/\alpha]/(\alpha-1) \qquad (10)$$

El steady state del consumo evaluado en (10) es:

$$c = k^{\alpha} + (1-\delta)k - k \tag{11}$$

Finalmente, es muy importante recordar que los resultados de la linearización y las conclusiones respecto al comportamiento dinámico **SOLO** son válidas en la proximidad del steady state y ***NADA*** mas.

E- Optimización Econométrica: Método de Maximum Likelihood (MLE)

Definición de la Función de Probabilidad Conjunta entre Variables Aleatorias

Consideremos una población en donde la variable aleatoria "y" tiene la siguiente función de densidad:

$$f(y,\theta)$$

Es decir, como toda función de densidad, depende de un grupo de parámetros (θ). Si por ejemplo estamos analizando la función de densidad de una normal la misma depende de dos parámetros:

$$\mu \ , \sigma$$

Entonces, para el caso en que la variable aleatoria "y" se distribuya como una normal, la función densidad de "y" seria:

$$f(y,\mu,\sigma) = [e^{-[(y-\mu)^2]/(2\sigma^2)}]/\sigma\sqrt{2\pi}$$

Recordemos que la probabilidad conjunta de un grupo de variables independientes se define como la multiplicatoria de sus respectivas funciones de densidad. Es decir:

$$\text{Funcion de Probabilidad Conjunta}(y_1, y_2, ...y_n) = f(y_1) * f(y_2) * ....f(y_n) = \prod_{i=1}^{n} f(y_i)$$

Por ejemplo, si hay tres variables que se distribuyen independientemente de una función de distribución normal, su respectiva función de probabilidad conjunta es:

$$\text{Funcion de Probabilidad Conjunta}(y_1, y_2, y_3) = \prod_{i=1}^{3} [e^{-[(y_i-\mu)^2]/(2\sigma^2)}]/\sigma\sqrt{2\pi}$$

Para el caso de la normal, definamos a la función de probabilidad conjunta de la siguiente forma:

$$\text{Funcion de Probabilidad Conjunta} = g(y_1, y_2, ...y_n, \mu, \sigma)$$

Es decir, la función de probabilidad conjunta depende de dos cosas: las observaciones muestrales $(y_1, y_2, ...y_n)$ y los parámetros de la función de densidad (μ, σ). En general, estamos acostumbrados a definir una función de probabilidad conjunta para un grupo ***dado y conocido*** de parámetros (en el caso de la normal, (μ,σ)).

Sin embargo, en muchas ocasiones uno cuenta sólo con las observaciones muestrales: $(y_1, y_2, ...y_n)$***, y a partir de las mismas intenta estimar los parámetros de la función de distribución de la cual provienen.*** Es aquí en donde entramos en el terreno del método de estimación Maximum Likelihood (MLE, Maximum Likelihood Estimation).

Definición de Likelihood y Log-Likelihood Functions

Supongamos que cuento con un número de observaciones muestrales $(y_1, y_2, ...y_n)$ y supongo que todas provienen de una misma función de distribución normal. El problema es que no sé cuáles son los parámetros (μ,σ). Es por esta razón que dadas dos condiciones:

Condición 1: Muestra $(y_1, y_2, ...y_n)$;

Condición 2: Supongo que $(y_1, y_2, ...y_n)$ provienen de una normal con μ y σ desconocidos. Puedo definir una función de probabilidad conjunta con parámetros desconocidos, ***la likelihood function,*** y utilizar a las observaciones $(y_1, y_2, ...y_n)$ para que me permitan estimar (μ, σ). La pregunta natural que surge ahora es: qué criterio voy a utilizar para estimar (μ, σ) a partir de $(y_1, y_2, ...y_n)$. Es precisamente en este criterio en donde se encuentra la esencia del método MLE. La idea es encontrar el valor de los parámetros (μ, σ) que me maximicen la likelihood function:

$$Max_{\mu,\sigma} \prod_{i=1}^{n} [e^{-[(y_i-\mu)^2]/(2\sigma^2)}] / \sigma\sqrt{2\pi}$$

La intuición detrás de esta maximización es muy sencilla: dado el supuesto que las observaciones muestrales se distribuyen como una normal, quiero encontrar los parámetros de esa distribución que me hagan máxima la probabilidad de que los datos muestrales que observo provengan de esa distribución. Obvio, el supuesto fundamental acá, es la función de distribución que genera las observaciones muestrales. Recordemos que yo sólo observo un vector de observaciones muestrales $(y_1, y_2, ...y_n)$, pero nunca voy a saber si este vector provino de una normal, eso es algo que supongo. Si por ejemplo, los datos muestrales provinieron de una log-normal y yo trato de maximizar una función de distribución normal, los resultados de la estimaciones van a ser muy malos (probablemente, desvíos standard muy altos).

Vemos que hasta acá expliqué la lógica suponiendo que los datos provienen de una normal. En realidad, el método es muy general, y se puede suponer cualquier distribución de probabilidad subyacente. Antes de continuar es importante recordar que si se aplica una transformación monótona a una función determinada, el resultado de la maximización no cambia. Un ejemplo de transformación monótona es la transformación logarítmica, esto implica que:

$$Max_{\mu,\sigma} \prod_{i=1}^{n} \frac{e^{-[(y_i-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}} = Max_{\mu,\sigma} \ln[\prod_{i=1}^{n} \frac{e^{-[(y_i-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}] = Max_{\mu,\sigma} \sum_{i=1}^{n} \ln[\frac{e^{-[(y_i-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}]$$

Recordemos que:

$$\ln(y_1 * y_2 * ....y_n) = \ln(y_1) + \ln(y_2) + ...\ln(y_n)$$

Como vemos, la transformación logarítmica tiene dos ventajas: convierte a la multiplicación en una suma y al momento de optimizar y calcular derivadas, las derivadas logarítmicas son muy fáciles de calcular. Es por esa razón (y dado que el máximo no cambia) que siempre se trabaja con la transformación monótona de la ***likelihood function,*** la cual se denomina ***log-likelihood function*** (éste es un término extremadamente común en este tipo de literatura).

Estimadores Maximum Likelihood

1- Función de Distribución Normal

Dado que ya describimos la intuición general del método, vamos a estimar a (μ, σ) ***suponiendo*** que la muestra observada $(y_1, y_2, ...y_n)$ proviene de una ***misma función de distribución normal*** (cuando digo misma me refiero al hecho que todas las observaciones provienen de una normal con los mismos parámetros (μ, σ)). Vamos a ilustrar la aplicación del método con varios ejemplos. El primero es para el caso de la distribución normal. Cada variable normal tiene la siguiente función de densidad:

$$f(y_i, \mu, \sigma) = [e^{-[(y_i-\mu)^2]/(2\sigma^2)}] / \sigma\sqrt{2\pi}$$

El logaritmo de esta variable es:

$$\ln[\frac{e^{-[(y-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}] = \ln(e^{-[(y-\mu)^2]/(2\sigma^2)}) - \ln(\sigma\sqrt{2\pi}) = -\frac{[(y-\mu)^2]}{2\sigma^2} - \ln(\sigma\sqrt{2\pi})$$

Recordemos que:

$$\ln(x * y) = \ln(x) + \ln(y); \quad \ln(\frac{x}{y}) = \ln(x) - \ln(y); \quad \ln(e^x) = x\ln(e) = x$$

Ahora vamos a sacar la derivada de la siguiente expresión con respecto a (μ, σ^2). Por simplicidad, tratemos al parámetro (σ^2) como un parámetro sin exponente; $(\sigma^2 = K)$ y luego hacemos la retransformación.

$$-[(y-\mu)^2]/2K - \ln(K^{0.5}\sqrt{2\pi})$$

Derivada con respecto a (μ):

$$2(y-\mu)/2K = (y-\mu)/K$$

Derivada con respecto a K:

$$-[(y-\mu)^2]/2K - \ln(K^{0.5}\sqrt{2\pi}) = -[(y-\mu)^2 * K^{-1}]/2 - \ln(K^{0.5}\sqrt{2\pi}) \Rightarrow$$

$$\text{La derivada es:}$$

$$[(y-\mu)^2 K^{-2}]/2 - (1/K^{0.5}\sqrt{2\pi})0.5K^{-0.5}\sqrt{2\pi} = [(y-\mu)^2 K^{-2}]/2 - 0.5/K$$

Recordemos lo siguiente:

$$\text{Der. de } a^x = xa^{(x-1)}; \quad \text{Der. de } \ln(x) = 1/x; \quad \text{Der. de } \ln(\mathrm{a}^x) = (1/a^x)xa^{(x-1)}$$

Ya tenemos entonces las derivadas con respecto a los dos parámetros (μ, K). Ahora hay que recordar que para maximizar las derivadas primeras deben forzarse a ser cero. Además recordemos que hasta ahora sólo calculé las derivadas de:

$$\ln([e^{-[(y-\mu)^2]/(2\sigma^2)}]/\sigma\sqrt{2\pi})$$

Pero la ***Log-Likelihood*** es:

$$\sum_{i=1}^{n} \ln([e^{-[(y_i-\mu)^2]/(2\sigma^2)}]/\sigma\sqrt{2\pi})$$

Con lo cual, dado que todas estas variables tienen la misma forma funcional, lo que tenemos que hacer es sumar derivadas idénticas e igualar toda esa suma a cero:

Derivada con respecto a (μ):

$$2(y-\mu)/2K = (y-\mu)/K \Rightarrow \sum_{i=1}^{n}(y_i - u)/K = 0 \text{ (Condicion de primer orden)}$$

Derivada con respecto a K:

$$\sum_{i=1}^{n}[((y-\mu)^2)/2K^2] - (n0.5/K) = 0 \text{ (Condicion de primer orden)}$$

Como vemos, nos queda un sistema de dos ecuaciones con dos incógnitas, la solución del mismo nos da los siguientes valores:

$$\hat{\mu} = (\sum_{i=1}^{n} y_i)/n; \qquad \hat{\sigma}^2 = (\sum_{i=1}^{n}(y_i - \overline{y})^2)/n$$

Vemos que estos resultados son familiares. Sin embargo, vemos que la estimación para la varianza es sesgada (divide por n y no por $(n-1)$). Sin embrago, para muestras grandes es insignificante, lo cual lo convierte en un estimador consistente de la varianza.

<u>2- Función de Distribución Poisson</u>

Como vemos, introdujimos el concepto de Maximum Likelihood Estimation para una función de distribución un tanto complicada: la normal. Para familiarizarnos con el concepto, vamos a aplicar el método para una función de distribución mas sencilla: función de distribución Poisson:

$$f(y_i,\lambda) = e^{-\lambda}\lambda^{y_i} / y_i!$$

La Likelihood Function para esta distribución es:

$$f(y_1, y_2, y_3, ..., y_n, \lambda) = \prod_{i=1}^{n} e^{-\lambda}\lambda^{y_i} / y_i!$$

Recordemos que:

$$\ln(e^{-\lambda}\lambda^{y_i} / y_i!) = \ln(e^{-\lambda}) + \ln(\lambda^{y_i}) - \ln(y_i!) = -\lambda + y_i \ln(\lambda) - \ln(y_i!)$$

La Log-Likelihood Function para esta distribución es:

$$\sum_{i=1}^{n}(-\lambda + y_i \ln(\lambda) - \ln(y_i!)) = -n\lambda + \ln(\lambda)\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\ln(y_i!)$$

Las condición de primer orden para maximizar el valor de la Log-Likelihood function con respecto al parámetro λ es:

$$-n + (\sum_{i=1}^{n} y_i)/\lambda = 0 \Rightarrow \lambda = (\sum_{i=1}^{n} y_i)/n$$

<u>3- Función de Distribución Bernoulli</u>

Recordemos que la función de distribución para una Bernoulli es:

$$f(y_i, p) = p^{y_i}(1-p)^{(1-y_1)}$$

La Likelihood Function para esta distribución es:

$$f(y_1, y_2, y_3, ..., y_n, \lambda) = \prod_{i=1}^{n} p^{y_i}(1-p)^{(1-y_i)}$$

Recordemos que:

$$\ln(p^{y_i}(1-p)^{(1-y_i)}) = \ln(p^{y_i}) + \ln((1-p)^{(1-y_i)}) = y_i \ln(p) + (1-y_i)\ln(1-p)$$

La Log-Likelihood Function para esta distribución es:

$$\sum_{i=1}^{n}[y_i \ln(p) + (1-y_i)\ln(1-p)] = \ln(p)\sum_{i=1}^{n} y_i + \ln(1-p)\sum_{i=1}^{n}(1-y_i)$$

La condición de primer orden para maximizar el valor de la Log-Likelihood function con respecto al parámetro p es:

$$(\sum_{i=1}^{n} y_i)/p - (\sum_{i=1}^{n}(1-y_i))/(1-p)) = 0 \Rightarrow (\sum_{i=1}^{n} y_i)/p = (\sum_{i=1}^{n}(1-y_i))/(1-p) \Rightarrow$$

$$(\sum_{i=1}^{n} y_i)(1-p) = (\sum_{i=1}^{n}(1-y_i))p \Rightarrow (\sum_{i=1}^{n} y_i)/n = p$$

4- Varianza Asintotica: Teorema de Cramer Rao

Como vimos, hasta ahora solamente nos concentramos en maximizar la Log-Likelihood Function pero no nos preguntamos si los estimadores MLE eran "buenos" o "malos". Es decir, no analizamos las propiedades de estos estimadores. Antes de analizar las propiedades, analicemos el siguiente teorema:

La varianza de un estimador consistente del parámetro poblacional θ nunca es menor que:

$$[-E(\frac{\partial \ln^2(L(\theta))}{\partial \theta^2})]^{-1}$$

Es decir, la varianza de un estimador consistente jamás es menor que la negativa de la esperanza de la inversa de la derivada segunda de la Log-Likelihood Function. Obvio, aplicar este teorema no es sencillo porque muchas veces la derivada segunda de la Log-Likelihood es una función no lineal cuya esperanza matemática puede ser muy difícil de calcular. Es por esta razón, que generalmente a este lower bound para la varianza se lo calcula así:

$$-[(\frac{\partial \ln^2(L(\hat{\theta}))}{\partial \hat{\theta}^2})^{-1}]$$

Es decir, al lower bound teórico se lo aproxima evaluando la negativa de la matriz hessiana en el valor del parámetro estimado. Cuál es la intuición de esta última expresión. En realidad es muy sencillo, recordemos que en primer lugar como estamos maximizando, las derivadas segundas en el máximo se tornan negativas. Es por eso que la expresión tiene el sigo menos delante, dado que la varianza nunca puede ser negativa. En segundo lugar, recordemos que la derivada segunda en el óptimo mide la curvatura que la función alcanza en eso óptimo. Cuanto mas curva es la función en el óptimo mas 'fuerte" es el máximo encontrado en relación a sus contrincantes próximos (recordemos que como la expresión es la inversa de la negativa de la hessiana, la hessiana en realidad esta dividiendo).

Si por ejemplo la matriz hessiana mostrase muy poca curvatura, el divisor seria relativamente pequeño y en consecuencia el lower bound para la varianza muy alto. Esto esencialmente nos dice que si el máximo encontrado no es mucho mejor que sus contrincantes próximos (o lo que es lo mismo, la hessiana muestra poca curvatura), el máximo encontrado exhibe una varianza alta ya que los contrincantes próximos se parecen mucho a él y en otra muestra podrían hasta ser mejores.

Propiedades de los Estimadores Maximum Likelihood

Las propiedades de los estimadores MLE para muestras pequeñas son en general, desconocidas. Es por eso que el método funciona mejor cuanto mas grande sea el tamaño de la muestra. Para muestras grandes, las propiedades asintoticas son las siguientes:

1- Los estimadores MLE son ***CONSISTENTES,*** es decir, a medida que el tamaño de la muestra crece, los estimadores muestrales convergen al verdadero valor del parámetro poblacional;

2- Son asintoticamente eficientes dado que alcanzan el límite mínimo de varianza definido por el teorema de Cramer-Rao:

$$\text{Varianza Asintotica}: \mathrm{V}(\hat{\theta}_{\mathrm{MLE}}) = -[(\frac{\partial \ln^2 (L(\hat{\theta}))}{\partial \hat{\theta}^2})^{-1}]$$

La relevancia de esta propiedad es muy fuerte: en el límite, los estimadores MLE son los de menor varianza posible. No hay ningún otro que pueda tener menor varianza.

1- Se distribuyen asintoticamente como una normal:

$$\hat{\theta} \rightarrow N(\theta, -[(\frac{\partial \ln^2 (L(\hat{\theta}))}{\partial \hat{\theta}^2})^{-1}])$$

Es muy importante entender por qué estamos poniendo tanto énfasis en la varianza de los estimadores MLE. La razón es muy sencilla: una vez estimados los parámetros, siempre es necesario testear hipótesis. Es imposible testear una hipótesis si no tenemos la matriz de varianza de los estimadores.

<u>Cálculo de la Varianza de los Estimadores Maximum Likelihood</u>

<u>1- Función de Distribución Normal</u>

Recordemos que la varianza de los estimadores MLE se calcula a través de la hessiana, la cual es la matriz de derivadas segundas de la Log-Likelihood Function. Recordemos que las derivadas primeras para el caso de la normal son:

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n} \frac{(y_i - u)}{K} = \sum_{i=1}^{n} \frac{y_i}{K} - \sum_{i=1}^{n} \frac{\mu}{K}$$

$$\frac{\partial \ln^2 L}{\partial \mu^2} = \sum_{\mathrm{i=1}}^{\mathrm{n}} 0 + \sum_{i=1}^{n} -\frac{1}{K} = 0 - \frac{n}{K} = -\frac{n}{\sigma^2}$$

$$\frac{\partial \ln L}{\partial K} = [\sum_{i=1}^{n} \frac{[(y-\mu)^2]}{2K^2}] - (n\frac{0.5}{K}) = [\sum_{i=1}^{n} \frac{[(y-\mu)^2]}{2} K^{-2}] - (n 0.5 K^{-1})$$

$$\frac{\partial \ln^2 L}{\partial K^2} = [-2\sum_{i=1}^{n} \frac{[(y-\mu)^2]}{2K^3}] + (n\frac{0.5}{K^2}) \Rightarrow$$

Analicemos el primer término de la expresión:

$$-2\sum_{i=1}^{n} \frac{[(y-\mu)^2]}{2K^3}$$

Recordemos que asintoticamente:

$$[\sum_{i=1}^{n} (y\text{-}\mu)^2]/n \rightarrow \sigma^2 \Rightarrow \sum_{i=1}^{n} (y\text{-}\mu)^2 \rightarrow n\sigma^2$$

Además:

$$K=\sigma^2 \Rightarrow K^3=\sigma^6 \Rightarrow -\frac{1}{K^3}\sum_{i=1}^{n}(y-\mu)^2=-\frac{n\sigma^2}{\sigma^6}=-\frac{n}{\sigma^4}\Rightarrow -2\sum_{i=1}^{n}\frac{[(y-\mu)^2]}{2K^3}=-\frac{n}{\sigma^4}$$

Analicemos el segundo término de la expresión:

$$n0.5/(K^2)$$

Recordemos que:

$$K=\sigma^2 \Rightarrow K^2=\sigma^4 \Rightarrow (n0.5/(K^2))=n/2\sigma^4$$

$$\frac{\partial \ln^2 L}{\partial K^2}=\frac{\partial \ln^2 L}{\partial(\sigma^2)^2}=-\frac{n}{\sigma^4}+\frac{n}{2\sigma^4}=\frac{-2n+n}{2\sigma^4}=-\frac{n}{2\sigma^4}$$

Se puede demostrar además que las derivadas cruzadas son cero. Entonces la hessiana nos queda:

$$\begin{matrix} -n/\sigma^2 & 0 \\ 0 & -n/2\sigma^4 \end{matrix}$$

Recordemos que a la hessiana hay que sacarle la inversa (recordemos que la inversa de una matriz es el adjunto de la matriz traspuesta dividida por el determinante de la matriz) y multiplicarla por –1 y evaluarla en el valor de los estimadores muestrales, llegando finalmente a la varianza de (μ,σ^2):

$$\begin{matrix} \hat{\sigma}^2/n & 0 \\ 0 & 2\hat{\sigma}^4/n \end{matrix}$$

Recordemos que el Teorema Central del Límite dice que la varianza de la media ($\hat{\mu}$) es la varianza de la variable subyacente dividida por n. Esto es exactamente lo que nos da el primer elemento de la matriz precedente. Con lo cual, la matriz de varianza covarianza estimada es la expuesta precedentemente.

2- Función de Distribución Poisson

Recordemos que las derivada primera para el caso de la Poisson es:

$$\frac{\partial \ln L}{\partial \lambda}=-n+(\sum_{i=1}^{n}y_i)/\lambda=-n+(\lambda^{-1})(\sum_{i=1}^{n}y_i)$$

Entonces:

$$\frac{\partial \ln^2 L}{\partial \lambda^2}=(-\lambda^{-2})(\sum_{i=1}^{n}y_i)=-(\sum_{i=1}^{n}y_i)/(\lambda^2)$$

$$\text{Si la evaluamos en el valor del estimador}: \hat{\lambda}$$

$$\text{Recordemos que}: \hat{\lambda}=(\sum_{i=1}^{n}y_i)/n \Rightarrow \hat{\lambda}n=(\sum_{i=1}^{n}y_i)$$

$$\frac{\partial \ln^2 L}{\partial \lambda^2}=-\hat{\lambda}n/\hat{\lambda}^2=-n/\hat{\lambda}$$

Recordemos que:

$$\text{Varianza Asintotica} : \mathrm{V}(\hat{\theta}_{\text{MLE}}) = -[(\frac{\partial \ln^2(L(\hat{\theta}))}{\partial \hat{\theta}^2})^{-1}] \Rightarrow -[(-n/\hat{\lambda})^{-1}] = \hat{\lambda}/n$$

Otra vez, recordando el Teorema Central del Límite, sabemos que la media muestral ($\hat{\lambda}$) tiene una varianza igual a σ^2 / n . Como en el caso de la Poisson $\sigma^2 = \lambda$, esto implica que la media muestral ($\hat{\lambda}$) tiene una varianza igual a $\hat{\lambda} / n$.

3- Función de Distribución Bernoulli

Recordemos que las derivada primera para el caso de la Bernoulli es:

$$\frac{\partial \ln L}{\partial p} = (\sum_{i=1}^{n} y_i) / p - (\sum_{i=1}^{n} (1-y_i)) / (1-p)) = [\sum_{i=1}^{n} y_i] p^{-1} - [\sum_{i=1}^{n} (1-y_i)](1-p)^{-1}$$

Entonces:

$$\frac{\partial \ln^2 L}{\partial p^2} = -[\sum_{i=1}^{n} y_i] p^{-2} - [\sum_{i=1}^{n} (1-y_i)](1-p)^{-2}$$

Recordemos que:

$$1)\ [\sum_{i=1}^{n} y_i] / n = p \Rightarrow np = [\sum_{i=1}^{n} y_i]$$

$$2) \sum_{i=1}^{n} (1-y_i) = \sum_{i=1}^{n} (1) - [\sum_{i=1}^{n} y_i] = n - np = n(1-p)$$

$$\frac{\partial \ln^2 L}{\partial p^2} = -[np]p^{-2} - [n(1-p)](1-p)^{-2} = -np^{-1} - n(1-p)^{-1} = -\frac{n}{p} - \frac{n}{1-p}$$

Sacando factor común:

$$\frac{-n(1-p)-np}{p(1\text{-}p)} = -\frac{n}{p(1-p)} \Rightarrow \frac{\partial \ln^2 L}{\partial p^2} = -\frac{n}{p(1-p)}$$

Recordemos que:

$$\text{Varianza Asintotica} : \mathrm{V}(\hat{\theta}_{\text{MLE}}) = -[(\frac{\partial \ln^2(L(\hat{\theta}))}{\partial \hat{\theta}^2})^{-1}] \Rightarrow -[(-\frac{n}{p(1-p)})^{-1}] = \frac{p(1-p)}{n}$$

Recordemos otra vez que al ser la varianza de la Bernoulli $p(1-p)$, nos queda que la media $\hat{p}$ tiene una varianza igual a σ^2 / n , o sea, $p(1-p)/n$.

Como vimos en los tres ejemplos, los estimadores MLE son consistentes con el Teorema Central del Límite que dice que la media muestral $\hat{x}$, en nuestro caso ($\hat{\mu}, \hat{\lambda}, \hat{p}$, respectivamente) exhibe una varianza igual a σ^2 / n .

Testeo de Hipótesis en Maximum Likelihood

1- Likelihood Ratio Test

La intuición detrás de un likelihood ratio test es muy sencilla. La idea es imponer una restricción en un (o unos) parámetro determinado. Una vez impuesta la restricción, maximizamos la "log-likelihood restringida". Obviamente, el valor de una función maximizada con restricciones es ***siempre*** menor al de la misma función maximizada sin restricciones. Cuanto mas fuerte sea la restricción, mayor será la diferencia entre la función maximizada restringida y la función maximizada sin restringir.

Para el caso específico de la estimación MLE, si la restricción es valida, imponerla no debería generar una reducción significativa en la log-likelihood function. Por lo tanto, el likelihood ratio test se basa en la diferencia entre la "unrestricted log-likelihood" y la "restricted log-likelihood". Definamos los siguientes conceptos:

$$L_U \text{ : Unrestricted Likelihood} \Rightarrow \ln(L_U) \text{ : Unrestricted Log - Likelihood}$$

$$L_R \text{ : Restricted Likelihood} \Rightarrow \ln(L_R) \text{ : Restricted Log - Likelihood}$$

$$L_R <= L_U \text{ , porque un optimo restringido nunca es superior a uno sin restringir}$$

Definamos al likelihood ratio test de la siguiente forma:

$$\lambda = L_R / L_U$$

Esta función debe estar entre cero y uno. Si λ es muy pequeño, esto implica que la restricción no es verdadera, es decir, no se condice con los datos de la muestra. La siguiente variable:

$$-2\ln(\lambda) = -2\ln(L_R / L_U) = -2[\ln(L_R) - \ln(L_U)] = 2[\ln(L_U) - \ln(L_R)]$$

Se distribuye como una chi-cuadrado, con grados de libertad igual a la cantidad de restricciones impuestas. Vamos entonces ahora, a testear la restricción de que:

$$\mu = 0.05$$

Recordemos, que siempre simulamos los datos con un valor:

$$\mu = 0$$

Entonces, debemos realizar las siguientes dos optimizaciones:

$$\text{Restringida} : Max_{\sigma} \sum_{i=1}^{n} \ln[e^{-[(y_i - 0.05)^2]/(2\sigma^2)} / \sigma\sqrt{2\pi}] = \ln(L_R)$$

Vemos aquí, que sólo maximizo con respecto a sigma, dado que la media se supone igual a 0.05. La segunda maximización, es la misma que la analizada anteriormente:

$$Max_{\mu,\sigma} \sum_{i=1}^{n} \ln[e^{-[(y_i - \mu)^2]/(2\sigma^2)} / \sigma\sqrt{2\pi}] = \ln(L_U)$$

Vemos aquí que maximizo con respecto a los dos parámetros, media y sigma. Entonces en este caso $2[\ln(L_U) - \ln(L_R)]$ se distribuye como una chi cuadrado con un grado de libertad.

Vamos entonces ahora, a testear la restricción simultánea de que:

$$\mu = 0.05, \sigma = 1.05$$

Recordemos, que siempre simulamos los datos con un valor:

$$\mu = 0, \sigma = 1$$

Entonces, debemos evaluar a la log-likelihood restringida:

$$\text{Restringida} : \sum_{i=1}^{n} \ln[e^{-[(y_i - 0.05)^2]/(2(1.05)^2)} / 1.05\sqrt{2\pi}] = \ln(L_R)$$

La maximización, es la misma que la analizada anteriormente:

$$Max_{\mu,\sigma} \sum_{i=1}^{n} \ln[e^{-[(y_i-\mu)^2]/(2\sigma^2)} / \sigma\sqrt{2\pi}] = \ln(L_U)$$

Vemos aquí que maximizo con respecto a los dos parámetros, media y sigma. Entonces $2[\ln(L_U) - \ln(L_R)]$ se distribuye como una chi cuadrado con dos grados de libertad.

2- Mixed Distributions

Hasta ahora sólo analizamos la aplicación de MLE para casos relativamente triviales, es decir, casos en donde la estimación pudo haberse hecho perfectamente mediante métodos lineales. Por ejemplo, para el caso de la normal, tanto su media y varianza puede estimarse utilizando estimadores standards. Lo mismo ocurre para la media y varianza de las distribuciones Poisson y Bernoulli.

Sin embargo, la verdadera utilidad de MLE es para casos de funciones no lineales, para las que los métodos tradicionales de estimación no ofrecen solución. Vamos a ver varios de estos casos, para los que en realidad se creó el método de Maximum Likelihood.

Supongamos que tenemos una distribución mixta:

$$\text{Funcion de Densidad de } y_i = \lambda[\frac{e^{-[(y_i-\mu_1)^2]/(2\sigma_1^2)}}{\sigma_1\sqrt{2\pi}}] + (1-\lambda)[\frac{e^{-[(y_i-\mu_2)^2]/(2\sigma_2^2)}}{\sigma_2\sqrt{2\pi}}]$$

Es decir, la variable aleatoria y_i, proviene de una distribución normal con media: μ_1 y varianza σ_1 con probabilidad λ y de una distribución normal con media: μ_2 y varianza σ_2 con probabilidad $(1-\lambda)$. Es como si fuera una lotería, yo no sé exactamente de cuál de las dos distribuciones proviene la observación y_i. Como vemos, si suponemos que sólo sabemos que la variable aleatoria y_i proviene de una de dos distribuciones normales, podemos estimar los 5 parámetros de la función ($\mu_1, \sigma_1, \mu_2, \sigma_2, \lambda$) utilizando MLE. Vemos que en este caso sería imposible aplicar algún método lineal de estimación.

$$Max_{\mu_1,\sigma_1,\mu_2,\sigma_2,\lambda} \sum_{i=1}^{n} \ln [\lambda(\frac{e^{-[(y_i-\mu_1)^2]/(2\sigma_1^2)}}{\sigma_1\sqrt{2\pi}}) + (1-\lambda)(\frac{e^{-[(y_i-\mu_2)^2]/(2\sigma_2^2)}}{\sigma_2\sqrt{2\pi}})]$$

Vemos entonces que podemos maximizar la log-likelihood con respecto a los 5 parámetros. Sin embrago, esta no es una tarea obvia dado que de acuerdo a la relación que exista entre ($\mu_1, \sigma_1, \mu_2, \sigma_2$) mejor será la información que le proveeremos a MLE.

Por ejemplo, si (μ_1, μ_2) son muy parecidas, al método le costará mucho distinguir de cuál de las dos distribuciones proviene una observación particular. En estos casos los resultados de la estimación van a ser muy malos. Por el contrario, si (μ_1, μ_2) son muy distintas, al método le será fácil calibrar una log-likelihood consistente con este hecho.

3- Probit. Diferencia entre una Función de Densidad y Distribución.

Una de las mayores ventajas de la aplicación de MLE se evidencia ***cuando la variable dependiente no es continua***. Recordemos que uno de los supuestos del modelo de Gauss-Markov (mínimos cuadrados ordinarios) es que la variable dependiente "y" es continua, es decir, puede asumir cualquier valor entre menos infinito y mas infinito.

Existen muchas situaciones en donde la variable dependiente "y" no es continua. Por ejemplo, supongamos que estamos analizando si las personas escogen irse de vacaciones o no. En este caso, dentro de la muestra vamos a tener a quienes se van de vacaciones (asignemos un uno por ejemplo a esto) y a quienes no se van de vacaciones (asignemos un cero a esto). Vemos entonces que la variable dependiente se comporta en forma dicotoma, es decir, asume solamente dos valores: 1, 0. Evidentemente, este comportamiento de la variable dependiente es muy diferente al supuesto por mínimos cuadrados, con lo cual si uno escoge estimar esto por mínimos cuadrados, el modelo va a estimar muy mal ya que los datos reales se comportan muy diferentemente respecto de los supuestos detrás del modelo lineal de mínimos cuadrados.

Es en este caso en particular en donde la estimación **NO LINEAL** se hace especialmente útil. Vamos a ver cómo podemos aplicar MLE a este caso. El modelo supone que existe una variable subyacente que no es observada y cuyo valor determina si una persona determinada escoge ir de vacaciones o no. Llamemos a esta variable subyacente y^*. Además, supongamos que esta variable subyacente depende de una serie de variables independientes, es decir, exógenas al modelo:

$$y^* = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Como siempre, dado que nunca podremos saber con total certidumbre qué es lo que generó la elección de una alternativa en particular, ponemos un término de error, ε, que como siempre, se distribuye como una normal con media cero y desvío standard 1.

$$\varepsilon \sim N(0,1)$$

Entonces, suponemos que la decisión de irse ($y = 1$) o no ($y = 0$) de vacaciones depende de lo siguiente:

$$y = 1 \text{ Si } y^* > 0; \qquad y = 0 \text{ Si } y^* \leq 0$$

Para estimar esto utilizamos un enfoque probabilistico:

$$\text{Prob(y = 1)} = \text{Prob}(y^* > 0) = \text{Prob}(\alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon > 0) = \text{Prob}[\varepsilon > -(\alpha + \beta_1 X_1 + \beta_2 X_2)]$$

Dado que la función de densidad normal es simétrica esto implica que:

$$\text{Prob}[\varepsilon > -(\alpha + \beta_1 X_1 + \beta_2 X_2)] = \text{Prob}[\varepsilon < (\alpha + \beta_1 X_1 + \beta_2 X_2)]$$

La expresión de la derecha es la función de probabilidad acumulada de una normal (o función de distribución, en contraposición a la función de densidad). Recordemos que la relación entre una función de densidad y una función de distribución para el caso de la normal es la siguiente:

$$\int_{-\infty}^{y_1} e^{-[(y_1-\mu)^2]/(2\sigma^2)} / \sigma\sqrt{2\pi}$$

Con $y_1 = \alpha + \beta_1 X_1 + \beta_2 X_2$. Es decir, la función de densidad mide cuál es la probabilidad de que el número y_1 salga. Mientras que la función de distribución mide cuál es la probabilidad de que el número que salga sea menor o igual a y_1. Es por esta razón que la función de distribución es la suma de todas las probabilidades asociadas a valores de la variable estocástica menores o iguales a y_1. Recordemos, que para el caso de variables continuas, la suma es el área o lo que es lo mismo la integral de la función de densidad. Es por esta razón, que la función de probabilidad acumulada (o función de distribución) de una normal se define así:

$$\int_{-\infty}^{y_1} \frac{e^{-[(y_1-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

Ahora bien, vamos a relacionar el concepto de probabilidad acumulada con el del Probit:

$$\text{Prob}[\varepsilon > -(\alpha+\beta_1 X_1+\beta_2 X_2)] = \text{Prob}[\varepsilon < (\alpha+\beta_1 X_1+\beta_2 X_2)] = \int_{-\infty}^{\alpha+\beta_1 X_1+\beta_2 X_2} \frac{e^{-[(\alpha+\beta_1 X_1+\beta_2 X_2-\mu)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

Dado que suponemos que $\mu = 0; \sigma = 1$, la expresión anterior nos queda:

$$\text{Prob}[\varepsilon < (\alpha+\beta_1 X_1+\beta_2 X_2)] = \int_{-\infty}^{\alpha+\beta_1 X_1+\beta_2 X_2} \frac{e^{-[(\alpha+\beta_1 X_1+\beta_2 X_2)^2]/2}}{\sqrt{2\pi}} = F(\alpha+\beta_1 X_1+\beta_2 X_2)$$

<u>Conclusión:</u> El modelo econométrico "Probit" supone que la probabilidad de que me vaya de vacaciones es igual a la función de probabilidad acumulada de la normal evaluada en las variables independientes.

Si por ejemplo, suponemos que las variables independientes son el ingreso y la edad, el Probit supone que irme de vacaciones está determinado por la función de probabilidad acumulada de la normal evaluada en los valores individuales de ingreso y edad. Supongamos que el individuo "i", tiene un ingreso de 100 pesos ($X_1^i = 100$) y una edad de 25 años ($X_2^i = 25$). Entonces, para el individuo "i", si escoge irse de vacaciones (y=1) la función de probabilidad acumulada es:

$$\text{Prob(y = 1)} = \int_{-\infty}^{\alpha+\beta_1 100+\beta_2 25} \frac{e^{-[(\alpha+\beta_1 100+\beta_2 25)^2]/2}}{\sqrt{2\pi}} = F(\alpha+\beta_1 100+\beta_2 25)$$

Si el individuo "i" escoge no irse de vacaciones (y=0) la función de probabilidad acumulada es:

$$\text{Prob(y = 0)} = [1 - \int_{-\infty}^{\alpha+\beta_1 100+\beta_2 25} \frac{e^{-[(\alpha+\beta_1 100+\beta_2 25)^2]/2}}{\sqrt{2\pi}}] = 1 - F(\alpha+\beta_1 100+\beta_2 25)$$

Obviamente, recordemos que tenemos muchos individuos en la muestra y que el concepto de likelihood function se refiere a la probabilidad conjunta de todas la observaciones. Definamos entonces la probabilidad conjunta:

$$\text{Likelihood Function} = \prod_{y_i=1} y_i[F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i)] * \prod_{y_i=0} (1-y_i)[1-F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i)]$$

Es esta función de probabilidad acumulada conjunta la que debemos maximizar con respecto a los parámetros α, β_1, β_2. Es importante aclarar que como vemos en la expresan anterior, hay dos multiplicatorias dentro de la likelihood function. La razón es que para cada individuo dentro de la muestra, caer en la primera o segunda multiplicatoria depende de si fue de vacaciones (y=1, primer multiplicatoria) o si no fue de vacaciones (y=0, segunda multiplicatoria). Vemos que si y_i es 1, $(1-y_i)$ es cero y viceversa. Es por esa razón que las dos multiplicatorias están acompañadas por los términos y_i y $(1-y_i)$ y solo una de ellas se aplica para cada observación.

Como siempre, por un problema de escala, transformamos esto en una log-likelihood:

$$\text{Log - Lik Function} = \sum_{y_i=1} y_i[\ln(F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i))] + \sum_{y_i=0} (1-y_i)[\ln(1-F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i))]$$

El método de MLE, maximiza la log-likelihood function con respecto a los tres parámetros:

$$\text{Max}_{\alpha,\beta_1,\beta_2} = \sum_{y_i=1} y_i[\ln(F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i))] + \sum_{y_i=0}(1-y_i)[\ln(1-F(\alpha+\beta_1 X_1^i+\beta_2 X_2^i))]$$

4- Probit. Interpretación de los Parámetros Estimados

Una vez estimados los parámetros, uno puede predecir la probabilidad de que una persona determinada escoja irse de vacaciones dados su ingreso y edad. Si por ejemplo los parámetros estimados son: $\alpha = 0.25, \beta_1 = 0.10, \beta_2 = 0.35$, el ingreso del individuo "i" es 100 pesos y su edad es 25 años, cargamos todo eso en la función de probabilidad acumulada y esto nos va a dar un número, el cual se debe interpretar como la probabilidad de que el individuo "i" escoja irse de vacaciones:

$$\text{Prob}(y=1) = [\int_{-\infty}^{0.25+0.10*100+0.35*25} \frac{e^{-[(0.25+0.10*100+0.35*25)^2]/2}}{\sqrt{2\pi}}] = \text{Numero}$$

Una computadora puede calcular fácilmente esta integral, con lo cual, la determinación de la probabilidad condicional es muy sencilla.

5- Ordered Probit

Este modelo es un poco mas complicado que el anterior y se usa cuando se tiene un poco mas de información que para el caso del Probit. Supongamos que en vez de tener una información dicotoma sobre la variable dependiente (0 o 1), tenemos (0,1,2,3….). Por ejemplo, este tipo de respuestas puede darse en dónde se va uno de vacaciones. El 0 indica Funes, el 1 Mar del Plata, el 2 Bariloche, el 3 Punta del Este, y el 4 Europa. Podemos modelar este proceso decisorio como dependiendo del ingreso del individuo:

$$\text{Ingreso} = \sum_{i=1}^{n} \beta_i X_i$$

Entonces, dado que no conocemos todo lo que determina la decisión, definamos a la variable subyacente como dependiendo del ingreso y algo desconocido (error):

$$y^* = \sum_{i=1}^{n} \beta_i X_i + \varepsilon_i$$

Entonces:

$$\begin{aligned}
&y = 0 \text{ si } y^* \le 0 \\
&y = 1 \text{ si } 0 < y^* \le u_1 \\
&y = 2 \text{ si } u_1 < y^* \le u_2 \\
&y = 3 \text{ si } u_2 < y^* \le u_3 \\
&y = J-1 \text{ si } u_{J-2} < y^* \le u_{J-1} \\
&y = J \text{ si } y^* > u_{J-1}
\end{aligned}$$

$$\begin{gathered}
\text{Prob}(0) = \text{P}(\beta\text{X} + \varepsilon \le 0) = \text{P}(\varepsilon < -\beta\text{X}) = \text{F}(-\beta\text{X}) \\
\text{P}(1) = \text{P}(\beta\text{X} + \varepsilon \le u_1 \;\&\; \beta\text{X} + \varepsilon > 0) = \text{F}(u_1 - \beta\text{X}) - \text{F}(-\beta\text{X}) \\
\text{P}(2) = \text{P}(\beta\text{X} + \varepsilon \le u_2 \;\&\; \beta\text{X} + \varepsilon > u_1) = \text{F}(u_2 - \beta\text{X}) - \text{F}(u_1 - \beta\text{X}) \\
\text{P}(J-1) = \text{P}(\beta\text{X} + \varepsilon \le u_{\text{J-1}} \;\&\; \beta\text{X} + \varepsilon > u_{J-2}) = \text{F}(u_{J-1} - \beta\text{X}) - \text{F}(u_{J-2} - \beta\text{X}) \\
\text{P}(J) = \text{P}(\beta\text{X} + \varepsilon > u_{J-1}) = \text{P}(\varepsilon > u_{\text{J-1}} - \beta\text{X}) = 1 - \text{F}(u_{J-1} - \beta\text{X})
\end{gathered}$$

Entonces la log-likelihood es:

$$\ln L = \sum_{y=0} \ln(\mathrm{F}(-\beta\mathrm{X})) + \sum_{y=1} \ln(\mathrm{F}(u_1 - \beta\mathrm{X}) - \mathrm{F}(-\beta\mathrm{X})) + ...\sum_{y=J} \ln(1 - \mathrm{F}(u_{J-1} - \beta\mathrm{X}))$$

Obviamente, igual que para el caso del Probit, cada individuo, de acuerdo a lo que haya escogido caerá en una de las tantas sumatorias de la log-likelihood. Finalmente, es importante aclarar que la optimización también se realiza en torno a los parámetros u_i los cuales indican a partir de qué nivel de ingreso los individuos escogen la siguiente alternativa. El calculo de los estimadores se hace en función de la siguiente optimización:

$$Max_{u_1,u_2,u_{J-1},\beta} = \sum_{y=0} \ln(\mathrm{F}(-\beta\mathrm{X})) + \sum_{y=1} \ln(\mathrm{F}(u_1 - \beta\mathrm{X}) - \mathrm{F}(-\beta\mathrm{X})) + ...\sum_{y=J} \ln(1 - \mathrm{F}(u_{J-1} - \beta\mathrm{X}))$$

<u>6- Estimación de Mínimos Cuadrados (OLS) utilizando Maximum Likelihood</u>

Recordemos que uno de los supuestos de OLS es que los errores se distribuyen como una normal con media cero y volatilidad sigma.

$$\varepsilon \sim N(0,\sigma^2)$$

Recordemos que $\varepsilon_i = y_i - \beta X_i$, entonces la probabilidad de que el error ε_i ocurra es:

$$f(\varepsilon_i, \mu, \sigma) = \frac{e^{-[(\varepsilon_i)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}} = \frac{e^{-[(y_i - \beta X_i)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

De esta forma podemos armar la likelihood y log-likelihood function:

$$\text{Likelihood Function} = \prod_{i=1}^{n} \frac{e^{-[(y_i - \beta X_i)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

$$\text{Log - Likelihood Function} = \sum_{i=1}^{n} \ln\left(\frac{e^{-[(y_i - \beta X_i)^2]/(2\sigma^2)}}{\sigma\sqrt{2\pi}}\right)$$

La log-likelihood se puede reexpresar así suponiendo que sólo hay un regresor para simplificar el álgebra:

$$\text{Log - Likelihood Function} = \sum_{i=1}^{n} \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{2\sigma^2}$$

$$\text{Log - Likelihood Function} = n\ln(1) - n\ln(\sigma) - n\ln(\sqrt{2\pi}) - \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{2\sigma^2}$$

$$\text{Log - Likelihood Function} = -n\ln(\sigma) - n\ln(\sqrt{2\pi}) - \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{2\sigma^2}$$

La FOC (first order condition) para Beta es:

$$\frac{\partial \ln \mathrm{L}}{\partial \beta} = \sum_{i=1}^{n} \frac{(y_i - \beta x_i)x_i}{\sigma^2} = 0 \Rightarrow \sum_{i=1}^{n} x_i(y_i - \beta x_i) = 0$$

Comparemos esta condición con la condición que Gauss le exigió a OLS. Recordemos que Gauss minimiza la suma de los errores al cuadrado:

$$Min_{\beta} \sum_{i=1}^{n} (y_i - \beta x_i)^2 \Rightarrow \mathrm{FOC} = \sum_{i=1}^{n} -2x_i(y_i - \beta x_i) = 0 \Rightarrow \sum_{i=1}^{n} x_i(y_i - \beta x_i) = 0$$

Como vemos, la condición de Gauss (OLS) es la misma que la que resulta de aplicar MLE. Es por esto que podemos sacar una conclusión fundamental: ***el valor estimado de Beta es idéntico tanto en OLS como en MLE.*** Vamos ahora a ver cuál es el valor estimado de sigma que resulta de aplicar MLE. La FOC (first order condition) para Sigma es:

$$\frac{\partial \ln \mathrm{L}}{\partial \sigma} = \sum_{i=1}^{n} -\frac{1}{\sigma} + \frac{(y_i - \beta x_i)^2}{\sigma^3} = 0 \Rightarrow \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{\sigma^3} = \frac{n}{\sigma}$$

$$\Rightarrow \sigma^2 = \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{n} = \sum_{i=1}^{n} \frac{(e_i)^2}{n}$$

Recordemos que en OLS el estimador de la varianza es:

$$\sigma^2 = \sum_{i=1}^{n} \frac{(e_i)^2}{n-k}$$

Es por esta razón que concluimos que OLS y MLE difieren en cuanto a la estimación de la varianza la cual es sesgada para el caso de MLE si bien en muestras grandes se hace insignificante.

7- Tests en Mínimos Cuadrados (OLS)

Antes de empezar definamos lo siguiente. Los estimadores OLS utilizando álgebra matricial se definen así:

$$b = (X^{'} X)^{-1} (X^{'} Y)$$

En donde b es un vector (n x 1). La varianza estimada del error term es:

$$e^i = y^i - \sum_{k=1}^{K} b_k x_k^i; \qquad \hat{\sigma}_\varepsilon^2 = (\sum_{i=1}^{N} (e^i)^2)/(N-K)$$

En donde K es el número total de parámetros incluida la constante. Los estimadores de la varianza de los parámetros estimados son:

$$\hat{\sigma}_b^2 = \hat{\sigma}_\varepsilon^2 (X^{'} X)^{-1}$$

Los t values de cada parámetro estimado son:

$$b_k / \hat{\sigma}_b$$

Existen diferentes formas de analizar el comportamiento de modelos econométricos. Una de estas formas es simular la variable dependiente como función de las variables independientes y estimar el modelo en función de esta información simulada. El objetivo de este tipo de análisis es analizar el "verdadero" comportamiento del modelo en un ambiente de laboratorio. Recordemos que los datos reales jamás se distribuyen exactamente como lo supone un modelo, este es un problema adicional de cualquier modelo que se base en supuestos.

La ventaja de simular datos y de esta forma crear condiciones de laboratorio es analizar cómo reacciona el modelo ante cambios diversos tales como: aumento en el tamaño de la muestra, cambios estructurales en los parámetros, contaminación en la información, etc.

8- Aumentos en el Tamaño de la Muestra

Existe una técnica muy conocida como "growing window technique". Esta técnica lo que hace es estimar varias veces el mismo modelo, pero para cada vez que se estima el modelo se le incrementa en una observación el tamaño de la muestra. Esta técnica esencialmente nos permite observar si los estimadores resultantes son sensibles a cambios en el tamaño de muestra. Recordemos que es deseable que los estimadores sean robustos es decir no sensibles a cambios en el tamaño de la muestra. Si por el contrario, tenemos un modelo cuyos parámetros estimados varían notablemente de estimación a estimación, esto quiere decir que el modelo que subyace a la regresión es altamente cambiante, lo cual viola el supuesto básico de todos estos modelos que suponen una misma forma estructural.

Es ahora en donde es muy útil simular y recrear las condiciones exactas en las que se basa OLS. Si las condiciones en las que se basa la simulación son coincidentes con ***TODOS*** los supuestos de OLS, aumentos en el tamaño de la muestra a través de un growing window, harán que los estimadores resultantes converjan a los verdaderos valores poblacionales (que conocemos dado que la información es simulada).

Veamos en primer lugar cómo se genera un growing window, este es el código respectivo:

Código

Límite Inferior del Growing Window
```
N=50;
```
Límite Superior del Growing Window
```
Lim=5000;
```
Generando Variables Independientes Hipotéticas
```
x=normrnd(0,1,Size,2);
```
Definiendo Parámetros Poblacionales Hipotéticos
```
Beta=[0.20;0.30;0.40;1];
```
Generando la Variable Dependientes Hipotética
```
y=(Beta(1,1))*ones(N,1)+(Beta(2,1))*x(:,1)+(Beta(3,1))*x(:,2)+normrnd(0,Beta(4,1),N,1);
```

Definiendo Acumuladores
```
Par=[0];
Par1=[0];
Par2=[0];
```

Comenzando el Growing Window
```
i=N;
while i<=Lim;
```
Incrementando Marginalmente las Variables Independientes Hipotéticas
```
x=[x;normrnd(0,1,1,2)];
```
Incrementando Marginalmente la Variable Dependiente Hipotética
```
y=(Beta(1,1))*ones(1,1)+(Beta(2,1))*x(i,1)+(Beta(3,1))*x(i,2)+normd(0,Beta(4,1),1,1);
y=[y;ay];
```
Agregando una Constante a la Matrix x
```
x=[ones(n(x,1),1) x];
```
Estimando los Parámetros de OLS
```
xols=inv((x')*((x)))*((x')*y);
```
Estimando la Varianza del Error Term (Beta(4,1))
```
Sigmaols=sqrt((sum((y-xols(1,1)*x(:,1)-xols(2,1)*x(:,2)-xols(3,1)*x(:,3)).^2))/(i-3));
```
Estimando la Varianza de los Estimadores
```
RegSig=(Sigmaols^2)*(inv((x')*((x))));
```
Eliminando la Constante de la Matriz X
```
x=x(:,2:3);
```

Calculando t values
```
tval=xols;
ii=1;
while ii<=n(xols,1);
         tval(ii,1)=xols(ii,1)/sqrt(RegSig(ii,ii));
         ii=ii+1;
end;
```

Incluyendo Resultados de la Estimación en el Acumulador de Resultados
```
Par=[Par;xols(:,1)'];
Par1=[Par1;tval(:,1)'];
Par2=[Par2;Sigmaols];
i=i+1;
```
Incremento de la Muestra en una Unidad
```
end;


Par=Par(2:end,:);
```

```
Par1=Par1(2:end,:);
Par2=Par2(2:end,:);
```

Gratificando los Estimated Betas

```
plot(Par);
AXIS([0 Inf 0 0.9])
title('Estimated Coefficient Dynamics as Sample N Increases')
```

Ahora que tenemos en claro cómo se genera un growing window, vamos a utilizar la técnica para analizar el comportamiento dinámico de OLS. Como el código precedente muestra, los datos fueron generados exactamente de acuerdo a los supuestos de Gauss-Markov. Por lo tanto sabemos que los estimadores OLS son consistentes.

Ahora que ya conocemos como se comportan los estimadores OLS cuando la simulación se condice perfectamente con los supuestos de Gauss-Markov, veamos qué pasa cuando estos supuestos se violan. En primer lugar, analicemos qué pasa cuando se genera ruido en la información, en este caso un incremento en el tamaño de la muestra agrega mas información pero también mas ruido. Hay muchas formas de generar ruido, una de ellas es hacer que los verdaderos parámetros poblacionales sean shoqueados constantemente por un error:

$$\text{Modelo de Acuerdo a Gauss - Markov } y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$\text{Modelo con Noise } y = \alpha + (\beta_1 + \varepsilon_1) X_1 + (\beta_2 + \varepsilon_2) X_2 + (\beta_3 + \varepsilon_3) X_3 + \varepsilon$$

$$\text{con } \varepsilon \sim N(0, \sigma^2)$$

$$\text{con } \varepsilon_1, \varepsilon_2, \varepsilon_3 \sim N(0,1)$$

Veamos en primer lugar, como cambia el código para agregar ruido:

```
while i<=Lim;
x=[x;normrnd(0,1,1,2)];

a1=normrnd(0,NoiseVolatility,1,1)*at;
a2=normrnd(0,NoiseVolatility,1,1)*at;
a3=normrnd(0,NoiseVolatility,1,1)*at;

y=(Beta(1,1)+a1)*ones(1,1)+(Beta(2,1)+a2)*x(i,1)+(Beta(3,1)+a3)*x(i,2)+normrnd(0,1,1,1);
……………………………….
```

Como vemos, dado que los errores que ahora contaminan a los parámetros poblacionales tienen media cero, el modelo va a converger a los mismos pero mas lentamente. Con respecto al modelo anterior (el que es totalmente coincidente con los supuestos de Gauss-Markov) el modelo va a converger con un tamaño superior de muestra y tener una varianza estimada mas alta dado que ahora la fuente del ruido no es sólo ε sino además $\varepsilon_1, \varepsilon_2, \varepsilon_3$.

Ahora generemos otro apartamiento, generemos un cambio estructural en el modelo. Es decir, hasta un determinado tamaño de muestra, el modelo tendrá un grupo determinado de parámetros poblacionales, a partir de ese punto, los parámetros poblacionales serán otros.

$$\text{Modelo Anterior al Cambio Estructural } y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$\text{Modelo Posterior al Cambio Estructural } y = \alpha' + \beta_1' X_1 + \beta_2' X_2 + \beta_3' X_3 + \varepsilon$$

$$\varepsilon \sim N(0,1)$$

Veamos en primer lugar, como cambia el código para generar el cambio estructural:

```
while i<=Lim;
x=[x;normrnd(0,1,1,2)];
Lim2=750;
if i<=Lim2;
```

Model Before Structural Change

```
y=(Beta(1,1)+a1)*ones(1,1)+(Beta(2,1)+a2)*x(i,1)+(Beta(3,1)+a3)*x(i,2)+normrnd(0,1,1,1);
y=[y;ay];
elseif i>Lim2;
```

Model After Structural Change

```
Structural=2;
```

Changing Parameter Scale (2 rather than 1)

```
Beta=Beta*Structural;
y=(Beta(1,1))*ones(1,1)+(Beta(2,1))*x(i,1)+(Beta(3,1))*x(i,2)+normrnd(0,1,1,1);
y=[y;ay];
end;.........................................
```

Con respecto al primer modelo, la única diferencia es que este modelo va a registrar un gran salto cuando la muestra alcanza 750 que es un donde ocurre la ruptura estructural. Este es el único cambio que debería ocurrir.

Finalmente, analicemos el mismo cambio estructural, pero además aumentemos la varianza del error term de uno a 2. El nuevo modelo es:

Modelo Anterior al Cambio Estructural $y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$

$$\varepsilon \sim N(0,1)$$

Modelo Posterior al Cambio Estructural $y = \alpha' + \beta_1' X_1 + \beta_2' X_2 + \beta_3' X_3 + \varepsilon'$

$$\varepsilon' \sim N(0,2)$$

Veamos cómo cambia el código para generar el cambio estructural:

```
while i<=Lim;
x=[x;normrnd(0,1,1,2)];
Lim2=750;
if i<=Lim2;
```

Model Before Structural Change

```
y=(Beta(1,1)+a1)*ones(1,1)+(Beta(2,1)+a2)*x(i,1)+(Beta(3,1)+a3)*x(i,2)+normrnd(0,1,1,1);
y=[y;ay];
elseif i>Lim2;
```

Model After Structural Change

```
Structural=2;
```

Changing Parameter Scale (2 rather than 1)

```
Beta=Beta*Structural;
y=(Beta(1,1))*ones(1,1)+(Beta(2,1))*x(i,1)+(Beta(3,1))*x(i,2)+normrnd(0,Structural,1,1);
y=[y;ay];
end;.........................................
```