Escuela de Administración y Negocios

Magister en Finanzas

Efectos de la convexidad en los cálculos de pérdida esperada: El caso argentino

Autor: Martín Senuy

DNI/Pas: 32.990.793

Director de Trabajo Final de Graduación: Gabriel Basaluzzo

Buenos Aires, febrero 2019

Resumen

En este trabajo, el autor buscará sentar las bases generales para el entendimiento de un problema que atañe a las entidades crediticias de todo el mundo y desarrollar parte de este para el caso argentino.

Cualquier entidad financiera cuyo negocio o parte de este implique el préstamo de sumas monetarias o activos corre el riesgo de no recibir un repago o que el mismo sea parcial. Pero, la situación financiera individual de cada deudor no debería ser el único factor bajo consideración. Hay factores exógenos que incrementan la potencial pérdida para el prestamista y que deben ser tenidos en cuenta al momento de efectuar una estimación.

Las nuevas normativas internacionales (FASB, IFRS9, CECL, etc.) incorporan requisitos regulatorios y contables más exigentes en cada actualización propia que realizan, inspirando avances de este tipo para la estimación y el tratamiento del riesgo crediticio.

En particular, trataré el problema de no incluir la convexidad dentro del cálculo de la probabilidad de default (PD) y daré una breve explicación sobre los efectos de ignorar las correlaciones existentes entre los componentes del cálculo de la pérdida esperada.

Indice

CAPITULO 1
INTRODUCCIÓN

Las normativas regulatorias internacionales (FASB, IFRS9, CECL, etc.) adaptan sus requisitos cada vez que son actualizadas para que no se reiteren crisis financieras como la del 2007-2009. Esta crisis reveló importantes deficiencias en las regulaciones financieras del momento, en las prácticas de administración de riesgos a nivel de las compañías y, de manera más general, una limitada comprensión de la dinámica del riesgo y, la interacción del crédito y las condiciones macroeconómicas.

Uno de los problemas principales fue la poca importancia otorgada al riesgo sistémico por los marcos regulatorios de todo el mundo. Dentro del riesgo sistémico, una parte importante es el riesgo sistemático de crédito. Este riesgo es la parte no diversificable del riesgo de crédito ya que afecta a toda la industria. Por ejemplo, los capitales regulatorios tradicionales podrían estar subestimando el riesgo sistemático al despreciar el impacto de que todos los bancos sufran un shock negativo simultáneamente. Es necesario evaluar las condiciones del riesgo sistemático de crédito y sus orígenes, ya que los cambios en el mismo pueden producir una elevada volatilidad en las tasas de default y explicar defaults conjuntos observados.

1.1 Motivación

Las situaciones de default se encuentran correlacionadas entre sí en un momento determinado. El primer motivo es la exposición a factores de riesgo comunes, ya sean observables o no. Por ejemplo, todas las empresas de un país están sujetas al ciclo económico del mismo, a la política monetaria y fiscal, a los precios del mercado, etc. Por otro lado, negocios en común u otros vínculos contractuales pueden originar un efecto contagio de default. Es decir, el default de una firma puede debilitar a otras con las que se encuentra vinculada, generando así una dependencia aún mayor.

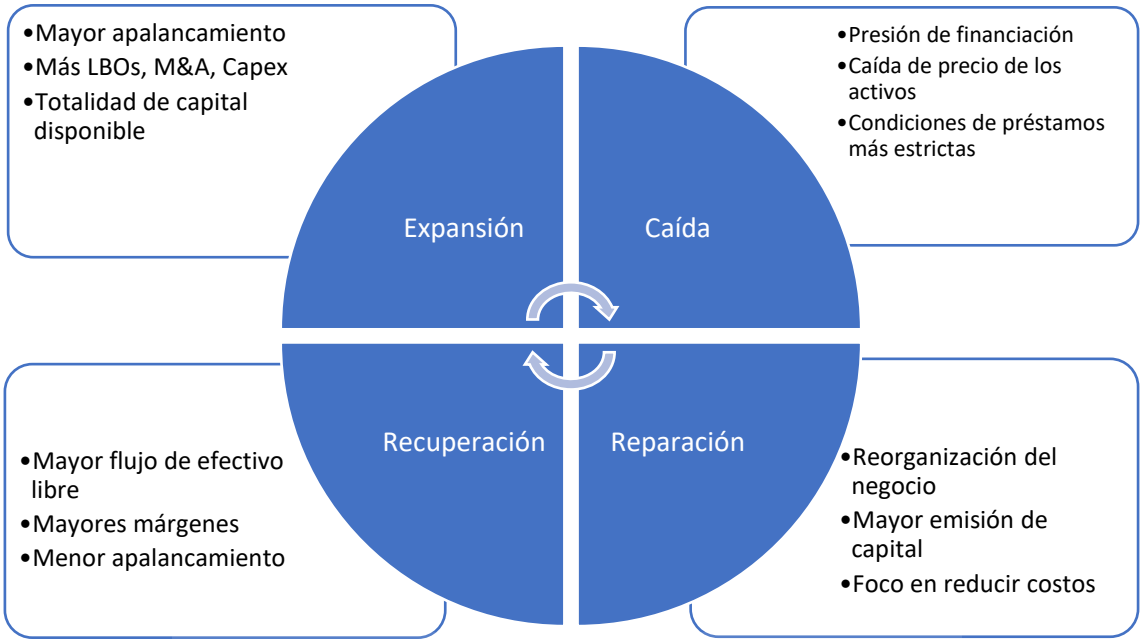
Es pertinente recordar que el interés inicial en los modelos de riesgo de crédito proviene de la necesidad de cuantificar la cantidad de capital económico necesario para cubrir la exposición total del banco a un potencial default. En este orden, el rol de un modelo de riesgo de crédito es tomar la información de las condiciones generales de la economía y las específicas de la firma y, obtener un spread de crédito, por sobre la tasa de interés a cobrar, que cubra la pérdida esperada (ECL). Dicha ECL se calcula a partir de tres componentes:

- Probabilidad de default (PD): Probabilidad de impago por parte del deudor;
- Exposición al Default (EAD): Porcentaje de la obligación que se encuentra expuesta a dicha falta de pago;
- Pérdida condicionada al Default (LGD por sus siglas en inglés): El porcentaje de la pérdida luego de descontar el recupero.

Aquí se encuentra la motivación del trabajo. En el mismo, mostraré la importancia de considerar la influencia del momento del ciclo de crédito, estimado a partir de variables macroeconómicas, al momento de calcular la PD y su efecto en el ECL. A esto me refiero con el impacto de considerar la convexidad de la PD. Asimismo, daré una breve explicación y un ejemplo sobre la correlación entre los tres factores que componen la ECL.

1.2 Ciclo del crédito y convexidad

En el gráfico a continuación, podemos ver un ejemplo de lo que llamamos ciclo del crédito. Si bien está reducido a cuatro etapas únicamente, es útil para comprender el concepto. Dado que puede haber más variables que lo caractericen y muchas combinaciones de las mismas, es necesario encontrar una convención o valor que las represente.



Gaurav, Ch. y Lawrence, R. en su trabajo *Convexity and Correlation Effects in Expected Credit Loss Calculations for IFRS9/CECL and Stress Testing (2016)*, utilizan un "Indice del ciclo del crédito". Ellos utilizan este índice para demostrar la convexidad (forma funcional no lineal) del modelo de la PD con respecto al riesgo sistemático. Es decir, mostrar que un shock sistemático negativo de crédito conlleva un cambio mucho mayor en las PDs comparado a un shock sistemático positivo de la misma magnitud. Estos efectos son severos en los modelos de PD, pero no tanto en los de EAD y LGD. Veamos un ejemplo con información de un único periodo, en el cual puede ocurrir un default, para lo que utilizan un factor de riesgo de crédito (Z) distribuido como una normal estandarizada.

Para el gráfico subsiguiente utilizan el modelo estándar de Vasicek para PD:

$$PD_{i,t} = \emptyset\left(-\frac{\emptyset^{-1}(p_{i,t}) + \sqrt{\rho}\cdot Z_t}{\sqrt{1-\rho}}\right)$$

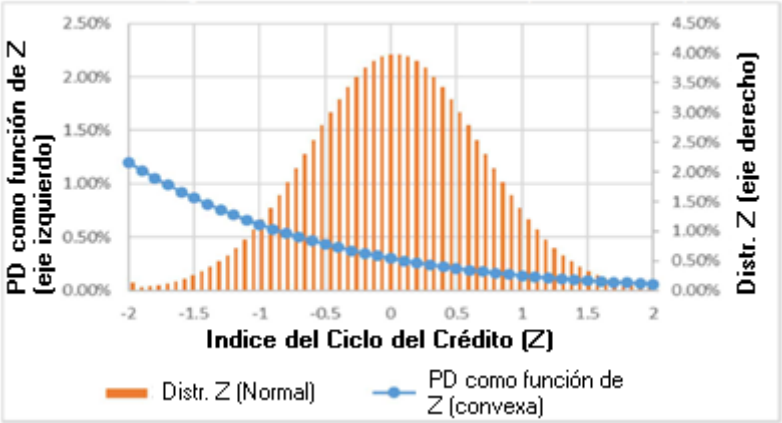
donde $PD_{i,t}$ es la PD para la empresa "i" en el momento "t"

$\emptyset$ - Función de distribución acumulativa de una normal estándar

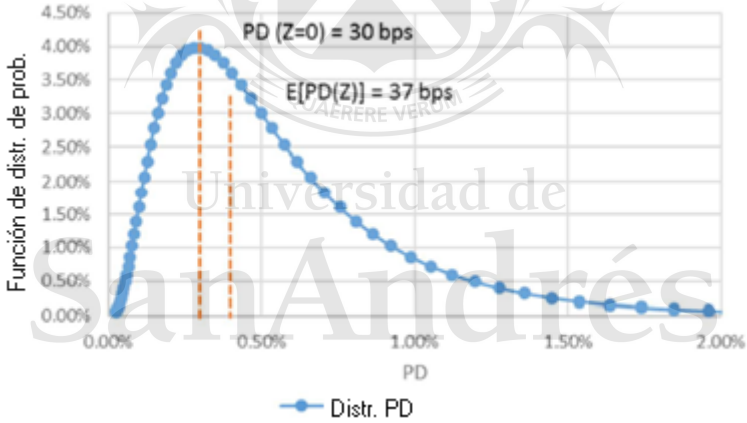
Z_t es el factor sistemático (índice del ciclo de crédito)

ρ – factor de correlación entre los activos de la empresa

Fig. 1: PD como función convexa de Z (datos ilustrativos)



Vemos claramente como la probabilidad de default es convexa teniendo una mayor pendiente hacia los valores negativos del índice Z que hacia los positivos. Subestimaríamos la PD en un momento de la economía donde la misma también influye en que los montos defaulteados sean mayores, conllevando previsiones inferiores a las necesarias para cubrir la exposición de la firma. Podemos ver un ejemplo en el siguiente gráfico:



Este gráfico nos muestra que si utilizáramos la PD sujeta al valor promedio del índice Z (0 por definición), obtendríamos una PD esperada menor que la esperanza de las PDs condicionadas a cada valor de Z. Una diferencia de 7 bps[1] puede parecer escasa, pero hay que recordar que esta debe multiplicarse por la exposición total de un banco (menos el valor de recupero).

Gracias a ambos gráficos, vemos el efecto que puede tener no considerar la convexidad a la hora de calcular la PD de una empresa.

[1] Bps es la sigla utilizada para "basis points" (puntos base) de tasas de interés. Cien (100) bps son 1%, por lo que 1 bps es un 0.01% de tasa.

1.3 Correlación

La convexidad no tiene un gran efecto en la EAD ni en la LGD, ya que dichas funciones no lineales no son suficientemente convexas cuando consideran el índice del ciclo de crédito (Z) como para crear desvíos sustanciales. Por otro lado, es necesario tener en cuenta el índice Z para abordar el problema de las correlaciones entre la PD, EAD y LGD. Idealmente, lo mejor sería utilizar un índice de este estilo ajustando al ciclo de cada industria-región, pero esto implica más trabajo y no siempre se encuentra disponible la información necesaria. Veamos la intuición de esta idea a partir de las siguientes fórmulas:

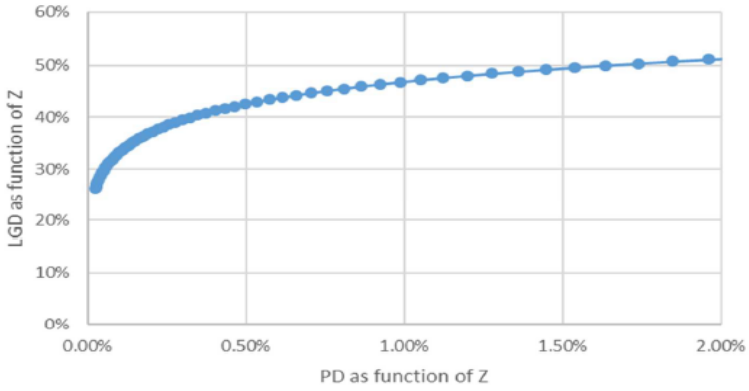
1) Ecuación incluyendo correlación y convexidad en base al índice:

$$ECL_{i,t} = E_s\left[PD_{i,s,t}(z_{s,t}) \cdot LGD_{i,s,t}(z_{s,t}) \cdot EAD_{i,s,t}(z_{s,t})\right]$$

2) Ecuación sin correlación pero con convexidad:

$$ECL_{i,t} = E_s\left[PD_{i,s,t}(z_{s,t})\right] \cdot E_s\left[LGD_{i,s,t}(z_{s,t})\right] \cdot E_s\left[EAD_{i,s,t}(z_{s,t})\right]$$

Como podemos notar, la diferencia entre ambas ecuaciones está en que la primera considera la interacción de las variables aleatorias, durante cada momento del ciclo, antes de calcular el valor esperado. Mientras tanto, la segunda ecuación considera el valor ponderado de cada componente ajustado por el índice del ciclo, pero de forma independiente. En el gráfico que se encuentra debajo, se puede observar un ejemplo de la correlación existente entre PD y LGD al incluir en el cálculo de ambas el índice de ciclo de crédito (Z).



Veremos un ejemplo numérico extraído del mismo artículo en la siguiente sección.

1.4 Ejemplo numérico

Es común considerar la exposición al default como un valor fijado para representar estos casos. Para los siguientes tres ejemplos de cálculo de ECL se utilizó la misma base de datos que para los gráficos anteriores. Estos cálculos consideran una PD = 0.03%, LGD = 39% y un valor fijo de EAD = $1.000.000:

1) Estimación sin convexidad ni correlación:

$$ECL(1) = PD(Z = 0) \cdot LGD(Z = 0) \cdot Fixed\ EAD = \$1.184$$

2) Estimación con convexidad sin correlación:

$$ECL(2) = E[PD(Z)] \cdot E[LGD(Z)] \cdot Fixed\ EAD = \$1.441$$

3) Estimación con convexidad y correlación:

$$ECL(3) = E[PD(Z) \cdot LGD(Z)] \cdot Fixed\ EAD = \$1.553$$

El incremento debido únicamente a la convexidad es de casi un 22%. Añadir la correlación luego supone un incremento aproximado de un 8% de (2) a (3). El aumento total es de más de un 30% y hay que tener en cuenta que, de haber incluido la correlación con la EAD, la diferencia hubiera sido aún mayor.

Dejemos que la siguiente tabla clarifique un poco más lo mencionado:

Z	Pr(Z)	PD	LGD	PD x LGD	EAD	ECL
<= -3	0.13%	2.21%	52%	1.14%	$1m	
-2.9	0.19%	2.08%	52%	1.07%	$1m	
-2.8	0.26%	1.96%	51%	1.00%	$1m	
.	.	.	.			
.	.	.	.			
.	.	.	.	.	.	
-0.1	3.94%	0.32%	40%	0.13%	$1m	
0	3.98%	0.30%	39%	0.12%	$1m	$ 1,185
0.1	3.98%	0.28%	39%	0.11%	$1m	
.	.	.	.	.	.	
.	.	.	.	.	.	
.	.	.	.	.	.	
2.8	0.07%	0.03%	27%	0.01%	$1m	
2.9	0.05%	0.03%	27%	0.01%	$1m	
> =3	0.13%	0.02%	26%	0.01%	$1m	
Prom ponderado		0.37%	39%		$1m	$ 1,442
Prom ponderado				0.16%		$ 1,554

Ignorando convexidad y correlación obtenemos un ECL más de un 30% inferior

Efecto significativo de la convexidad sobre la PD

Efecto no significativo de la convexidad sobre el LGD

Efecto significativo de la convexidad sobre el ECL

Efecto significativo de la correlación sobre el ECL

1.5 Descripción del trabajo

En los capítulos siguientes, abordaremos el objetivo del trabajo. El mismo es explicar una posible solución al cálculo de la probabilidad de default para incluir el efecto de la convexidad en el mismo. En el capítulo 2, explicaremos brevemente una metodología comúnmente utilizada por grandes bancos a nivel mundial. Más detalles de dicha metodología podrán verse en los Anexos o se encontrarán referencias a los trabajos completos en la bibliografía. En el capítulo 3, intentaremos aplicar dicha metodología al caso argentino. Se explicarán las variables elegidas para dicho propósito y se intentará brindar una solución numérica para las ecuaciones. Asimismo, explicaré los resultados y realizaré una comparación con el caso base. Por último, se brindarán las conclusiones sobre dichos resultados y, algunas observaciones sobre lo que creo que puede analizarse o mejorarse a futuro. El capítulo 4 incluirá las referencias bibliográficas utilizadas para el desarrollo del presente trabajo. En los anexos se encontrará material más detallado de ciertos componentes del trabajo realizado, ya sean modelos utilizados, tablas de datos, códigos de programación, etc.

CAPITULO 2
EVOLUCIÓN TEÓRICA

En el presente capítulo describiremos los métodos y/o modelos utilizados en el desarrollo del caso. Se incluirán los desarrollos que consideré indispensables para el entendimiento de dichos modelos. Un desarrollo más detallado o completo podrá ser visto en los Anexos o en las referencias citadas. Principalmente, explicaré como aplicar el modelo de Vasicek de valuación de portafolios de préstamos y, luego, como se complementa este con el Filtro de Kalman para ajustar las PD en un momento dado del ciclo del crédito. El Filtro de Kalman será explicado con la mayor simpleza posible con el propósito de comprender el presente trabajo.

2.1 Vasicek: Modelo de valuación de portafolios de préstamos

El trabajo de dicho autor fue publicado en el año 2002 en *Risk*. La motivación de aquel trabajo fue encontrar la distribución de pérdidas del portafolio con el fin de estimar el monto de capital necesario para sostener un portafolio de títulos de deuda. Por ejemplo, en un portafolio de préstamos precisamos determinado capital para soportar las pérdidas por los incumplimientos de pago. Su enfoque se basa en considerar que una firma incurre en default en caso de que sus activos sean menores que sus obligaciones al momento del pago de estas.

El modelo de Vasicek asume que, la evolución del valor de los activos de un determinado deudor viene dado por un efecto combinado de un factor sistemático y uno idiosincrático. El autor asume una estructura de default equi-correlacionada y gaussiana. Es decir, cada deudor *i* defaultea si una determinada variable aleatoria X_i cae por debajo de un umbral, y estas variables X_is son todas normales y equi-correlacionadas. El valor de los activos del deudor *i* en el momento *t* esta dado por[2]:

$$X_{it} = S_t\sqrt{\rho} + Z_{it}\sqrt{1-\rho}$$

[2] Ver justificación en Anexo A.

En esta ecuación, S representa el factor sistemático y, por eso, no depende del valor de *i,* ya que no depende de cada firma sino del momento en el tiempo. Por otro lado, Z es el factor idiosincrático correspondiente a cada empresa. Finalmente, ρ es la correlación entre los activos de dos deudores diferentes. Entendemos que esta correlación viene dada por el factor sistemático y, es por eso, que está multiplicada por el mismo, y (1- ρ) está multiplicada por el factor individual. Por último, los valores de Z son variables normales estándar independientes entre sí.

Adentrándonos en el modelo, vemos que usa tres datos para calcular la probabilidad de default de una clase de activos. El primero es la PD "a través del ciclo" (TTC PD) específica para esa clase. Esta PD es ni más ni menos que la PD original que uno conoce, que se considera constante a través del ciclo de crédito antes de hacerle ningún ajuste por convexidad. El siguiente dato es el factor común del portafolio o la industria como, por ejemplo, el ciclo del crédito que ya hemos mencionado o un índice que explique el estado de la economía a lo largo del periodo analizado. Llamaremos *S* a esta variable "estado". Por último, el tercer dato es la correlación entre los activos, ρ. Entonces, podemos concluir que el término $S_t\sqrt{\rho}$ es la exposición de la empresa al riesgo sistemático y que el término $Z_{it}\sqrt{(1-\rho)}$ representa el riesgo idiosincrático de la misma.

En el Anexo A se comprueba que, el deudor *i* defaultea en caso de $X_i < c$, siendo c el valor que determina la probabilidad incondicional de default (TTC PD):

$$\Pr(X_i < c_i) = N(c_i) = TTC\ PD_i = p^*$$

La probabilidad de default condicionada al estado *S* es la siguiente:

$$\Pr(X_i < c \mid S) = \Pr\left(S\sqrt{\rho} + Z_i\sqrt{1-\rho} < c \,\middle|\, S\right) = \Pr\left(\frac{c - S\sqrt{\rho}}{\sqrt{1-\rho}} \mid S\right)$$

Combinando ambas ecuaciones obtenemos:

$$\Pr(X_i < c \mid S) = N\left(\frac{c - S\sqrt{\rho}}{\sqrt{1-\rho}}\right) = N\left(\frac{N^{-1}(p^*) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

Para este modelo las correlaciones se consideran entre 0 y 1, $0 \leq \rho < 1$. Estas correlaciones, junto con el factor estocástico *S*, son los factores de los que depende la probabilidad de default de una empresa.

Cabe destacar que el factor *S* es a lo que me refería en el capítulo anterior con "Indice de ciclo de crédito" (Z). Si bien en este modelo se busca relacionar *S* a un escenario macroeconómico, es la misma idea ya que el ciclo de crédito viene dado por condiciones macroeconómicas. Además, la distinción que busca hacer Vasicek por clase de activos es la misma que mencioné anteriormente por industria o región. Idealmente, uno debería segmentar los préstamos en categorías lo más específicas posible para aislar los efectos homogéneos de cada categoría en particular. Pero, esta segmentación se puede realizar hasta donde la calidad y la cantidad de la información que se posee para trabajar lo permita. De nada vale establecer miles de categorías si luego las estimaciones son poco confiables. Por eso, muchas veces uno utiliza las mismas variables macroeconómicas para todas las categorías o usa pocas categorías de préstamos/empresas para realizar los ajustes a la PD. El objetivo es extraer, de estas variables observables, la información común a todos los individuos (préstamos, empresas, etc.) para representar la variable "estado" e implementarla en el cálculo de la PD ajustada. Para esto puede ser utilizado el filtro de Kalman. La ventaja de esta técnica es que permite que las variables "estado" sean magnitudes no observables. El filtro de Kalman será explicado en la siguiente sección.

Finalizando esta sección, haré un breve análisis de la última fórmula utilizada para describir la PD:

$$\Pr(X_i < c \mid S) = N\left(\frac{N^{-1}(p^*) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

Analizando los dos factores que dijimos que eran clave para la misma, el factor de correlación y la variable estado, podemos notar lo siguiente:

- En niveles normales de correlación y asumiendo que *S* es una variable normal estándar, como la variable Z de ciclo de crédito que hemos analizado antes, podemos ver que el valor de *S* es el que determina si la PD ajustada es mayor, menor o igual que la PD TTC.
- Respecto al punto anterior, cabe mencionar que se deben analizar las combinaciones posibles de valores de *S y* ρ. Si bien un escenario negativo pareciera incrementar de forma inequívoca la PD (siendo consistente con

lo expuesto) y un escenario positivo disminuirla, estas variaciones realmente dependen del nivel de dicha combinación de valores.

- Cuando $\rho = 0$, la PD es la misma que la TTC PD sin importar el valor de la variable *S.* Cada PD dependería únicamente del factor idiosincrático y las variables comunes, que generarían efectos a través de las correlaciones entre los activos, no tendrían ningún efecto en la misma.
- Cuando $\rho \to 1$, la PD tiende a 1 o a 0 dependiendo del signo y magnitud del factor *S.* Esto evidencia que, ante el caso de una alta correlación entre los activos, tanto los efectos positivos como negativos derivados de la variable "estado" pueden resultar extremos.

2.2 Filtro de Kalman

La idea principal de cualquier filtro es dar sentido a información ruidosa y, aclararla o separar el ruido de la información. Hay muchas razones para que la información sea ruidosa o no observable, como errores de medición o incongruencias en la frecuencia de medición. Estos problemas son de interés para las personas que buscan monitorear y controlar sistemas dinámicos. Es imprescindible lograr una buena medición de las condiciones del momento. Los métodos de modelación utilizados en sistemas dinámicos son conocidos como "Métodos de espacio de estado" (State space method) y la solución óptima de un sistema dinámico linear bajo condiciones gaussianas se obtiene a partir del Filtro de Kalman. Es decir, el filtro de Kalman lineal es el estimador óptimo, posee el menor error cuadrático alrededor de la media, en el caso de que las variables observadas y las perturbaciones sean conjuntamente gaussianas. Para otros casos, es el mejor pero solo entre los estimadores lineales.

Es un marco genérico usado para modelar un sistema dinámico, ya sea en tiempo discreto o continuo, en el cual uno puede separar lo que el usuario puede observar y el estado real de las variables de interés del sistema. En análisis de series temporales, uno intenta obtener el estado de un sistema dadas las mediciones de determinadas variables. Este estado del sistema representa varios componentes no observables como tendencias y estacionalidades.

En particular, el filtro de Kalman es una serie de ecuaciones que permite la actualización de un estimador a medida que se encuentran disponibles nuevas observaciones. Primero, forma una variable predictiva óptima del vector S de la variable de estado no observado dados sus valores calculados previamente. Es decir, extrapola estos componentes hacia el futuro. Estos cálculos de las variables de estado no observado, posteriormente, se actualizan utilizando la información proporcionada por las variables observadas.

Digamos que Y_t es una serie de "n" variables observables en un tiempo t, siendo $t = 0, 1, \dots, T$. Es decir, Y_t representa una serie temporal multivariada. El modelo puede describirse de la siguiente forma:

$$(1)\ Y_t = \beta s_t + \emptyset Y_{t-1} + \psi \varepsilon_{t-1} + \varepsilon_t$$

$$s_{t+1} = \gamma s_t + \eta_{t+1}$$

$$s_0 \sim N(S_0, p_0)$$

$$\varepsilon_t \sim N(0, I\sigma_\varepsilon^2)$$

$$\eta_t \sim N(0,1)$$

$$t = 0, 1, \dots, T$$

Donde s_t es el proceso estocástico escalar que representa el estado de las condiciones macrofinancieras; β es el vector de dimensión *m* de cargas del factor; $\emptyset$ es una matriz diagonal m × m de los coeficientes AR(1); ψ es una matriz diagonal de los coeficientes MA(1); γ es el coeficiente autorregresivo del proceso estocástico no observable s_t; ε_t es un proceso de ruido blanco multivariado gaussiano con varianza diagonal y matriz de covarianza $I\sigma_\varepsilon^2$; y η_t es un ruido blanco gaussiano escalar con varianza unitaria.

Para cerrar el modelo, debemos especificar las condiciones iniciales de s_t. Entonces, suponemos que s_0 está distribuido normalmente con media S_0 y varianza p_0.

El factor común S_t en el tiempo *t* se obtiene como valor esperado del proceso s_t. Los parámetros del modelo pueden calcularse reescribiendo la ecuación (1) en forma de espacio de estado y aplicando el filtro de Kalman. Una representación de estado de espacio equivalente de (1) puede obtenerse como:

$$(2) \qquad Y_t = Z\alpha_t$$

(3) $$\alpha_{t+1} = T\alpha_t + Ru_{t+1}$$
$$\alpha_0 \sim N(a_0, P_0)$$

La ecuación (2) es la ecuación de medición que relaciona las variables macroeconómicas y financieras observables con las variables de estado no observables α_t. La ecuación (3) es la ecuación de estado y consigna las dinámicas de la variable de estado latente α_t. El algoritmo funciona como un proceso de dos pasos. En el paso de predicción, el filtro de Kalman genera estimaciones de las variables de estado actual, así como la incertidumbre correspondiente de las estimaciones. Una vez que la siguiente medición se observa (con algo de ruido), estas estimaciones se actualizan.

Con el modelo escrito en forma de espacio de estado, la estimación de los parámetros puede obtenerse mediante el método de máxima verosimilitud.

El filtro de Kalman permite calcular la media condicional $E(\alpha_t|Y_t, Y_{t-1}, \dots, Y_0) = a_t$ y la varianza condicional $Var(\alpha_t|Y_t, Y_{t-1}, \dots, Y_0) = P_t$.

Las recursiones de Kalman[3] que nos permiten computar el factor común S_t como el primer elemento del vector a_t son:

(4) $$v_t = Y_t - Za_t$$
(5) $$F_t = ZP_tZ'$$
(6) $$K_t = TP_tZ'F_t^{-1}$$
(7) $$L_t = T - K_tZ$$
(8) $$a_{t+1} = Ta_t + K_tv_t$$
(9) $$P_{t+1} = TP_tL_t + Q$$

que se iteran desde *t = 0*. La matriz K_t en (7) se conoce como la ganancia de Kalman.

Las recursiones de Kalman se derivan de la fórmula de la media condicionada y de la varianza condicionada en una distribución multinormal. Una distribución normal es caracterizada plenamente por sus dos primeros momentos y la función de verosimilitud exacta se obtiene como subproducto del algoritmo del filtro de Kalman. Para inicializar el filtro, necesitamos especificar a_0 y P_0. Una opción es

[3] Puede verse un ejemplo práctico de las recursiones en el Anexo C.

considerar estos parámetros como fijos y estimarlos utilizando máxima verosimilitud. Asuma que $\alpha_t \sim N(a_0, P_0)$, donde a_0 y P_0 son conocidos, y ψ denota el conjunto de parámetros a estimar, es decir, $\psi = \{Z, T, R\}$, la función de verosimilitud puede escribirse como:

$$L(\psi) = p(Y_o) \prod_{t=1}^{T} p(y_t | Y_{t-1}) \tag{10}$$

Dado que:

$$E(y_t | Y_{t-1}) = E(Z\alpha_t | Y_{t-1}) = Za_0$$
$$v_t = y_t - Za_t$$
$$Var(y_t | Y_{t-1}) = Var(v_t) = H$$

Entonces:

$$L(\psi) = \prod_{t=1}^{T} p(v_t) \qquad v_t \sim N(0, H)$$

Por lo tanto, el valor de la función de verosimilitud puede computarse directamente a partir de las recursiones de Kalman en las ecuaciones (4)-(9) y, el máximo del $log(L(\psi))$ puede obtenerse numéricamente. Para más detalles de la derivación de la aplicación de máxima verosimilitud ver el Anexo B.

CAPITULO 3
EL CASO ARGENTINO

Este capítulo es el punto neurálgico del trabajo. Aquí, mencionaré ciertas pautas y consideraciones que suelen tenerse para el cálculo de la pérdida esperada (ECL) en la práctica real. Asimismo, realizaré un análisis del caso argentino y ejemplificaré la aplicación de este método en este contexto, bajo determinados supuestos y condiciones.

3.1 Consideraciones y supuestos para la aplicación

Los bancos y compañías financieras suelen asignar calificaciones a sus clientes o potenciales clientes para otorgarles préstamos. Dichas calificaciones pueden basarse en el valor de sus activos, su nivel de obligaciones, la rama del negocio a la que se dedican, volumen de transacciones, etc. Las mejores calificaciones son otorgadas a aquellos con una menor posibilidad de impago, pagando por ello una tasa menor en cuanto a spread de crédito. Pero además de la probabilidad de impago, cada calificación tiene asignada una probabilidad de transición hacia las otras calificaciones que componen la grilla. Dichas transiciones pueden ser hacia calificaciones superiores como inferiores. Todas estas probabilidades son ordenadas en una matriz llamada *Matriz de transición*.

Es importante considerar esta matriz al momento de realizar proyecciones, ya que es posible que una empresa que no incurra en default sí sufra una transición negativa (o positiva) en su calificación. Esto afectará a las futuras PD de la misma, por lo que al momento de realizar estimaciones a futuro de la pérdida esperada (ECL) hay que tener en cuenta cada posible transición para cada periodo. Dichas probabilidades son las que he llamado “a través del ciclo” (TTC) y se consideran permanentes durante determinado periodo, por lo que las variaciones que se puedan encontrar en la matriz se deberán a los ajustes realizados por el ciclo económico/crediticio.

Cabe destacar que estos ajustes sirven para el cálculo de las primas a cobrar y cubren únicamente las pérdidas esperadas. Pero las normativas también exigen determinada cobertura (reservas de capitales) para las pérdidas inesperadas que exceden a esta estimación hasta el valor que represente el 99.9% de la distribución acumulada de pérdidas. A este excedente se lo denomina “Pérdida inesperada” (UL – unexpected loss).

Algunas normativas piden incluir dentro de las pérdidas esperadas un adicional por “Incertidumbre de gran préstamo”, a fin de tener en cuenta que no hay uniformidad en los montos de los préstamos y considerar el default de uno que exceda el valor medio de los mismos. Este es un ajuste cuantitativo. También suelen hacerse ajustes cualitativos a criterio de la gerencia de riesgo, en base a su experiencia con determinados préstamos, líneas de negocios o clientes.

Es necesario entender que todos estos ajustes incrementan la pérdida esperada, pero, a su vez, disminuyen la pérdida inesperada. Al reconocer mayor riesgo a los préstamos, la incertidumbre que queda es menor. Esto se refleja en un menor capital que la compañía debe aprovisionar, por lo que es libre de disponerlo de la manera que crea mejor. Este beneficio, sumado al cobro de primas más altas explicado, acarrea una mejoría en los resultados de dichas empresas.

Para el caso bajo análisis, consideraré las matrices de transición como ya dadas y aplicaré las proyecciones del ajuste por el estado de la economía a los valores de esta que representen la probabilidad de default de la empresa analizada. Este ajuste es el mismo para cada valor de PD. Quedará para futuros análisis el hecho de evaluar si esto es correcto o si también hubiera que tener en cuenta un mayor ajuste sobre determinadas categorías.

3.2 Breve reseña histórica

En las últimas décadas, la economía argentina ha sido muy fluctuante. Etapas de recesión y crisis severas se han intercalado con periodos de crecimiento y recuperación. Dichas oscilaciones han generado un impacto sobre todos los sectores de la economía, y el sector crediticio no fue la excepción. El impacto que puede generar una economía cambiante e impredecible como lo es la

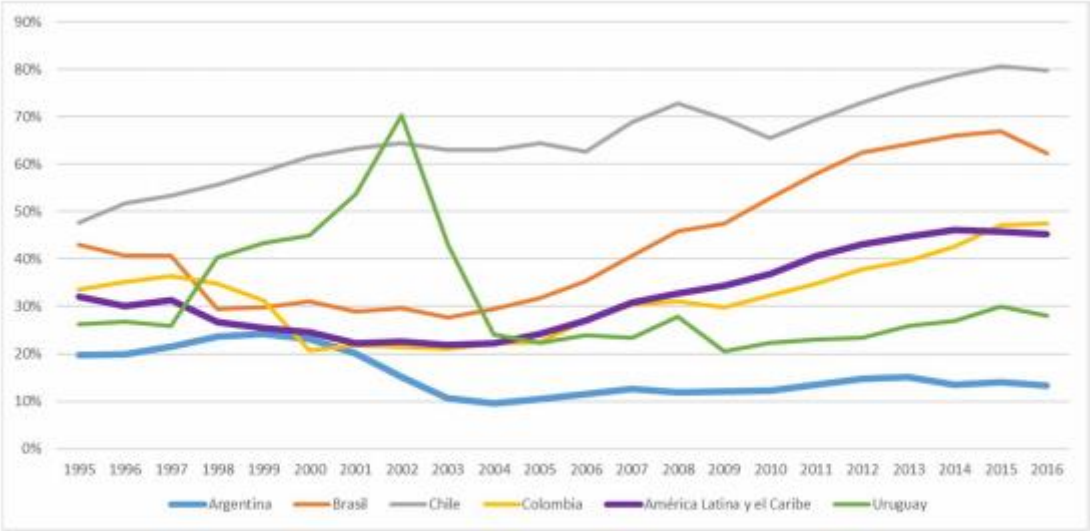
economía argentina sobre el sistema crediticio es de gran importancia ya que, una crisis financiera conlleva un derrame al resto de los sectores de la economía generando problemas severos, difíciles de solucionar en el corto plazo. Esto hace que el estudio del impacto de variables macroeconómicas en el sistema crediticio sea sumamente relevante, debido a que podría ser utilizado como un mecanismo de prevención frente a conflictos futuros.

La crisis económico-financiera del año 2001 fue un punto de inflexión para el sistema financiero argentino ya que quedo reducido al mínimo tamaño posible. En este contexto económico, el sistema bancario argentino se enfrentaba a una alta exposición al sector público, un fuerte descalce de plazos y de moneda. El resultado de las medidas adoptadas por el gobierno derivó en graves problemas en el sistema financiero en materia de solvencia y liquidez. Todo esto propició un gran nivel de default del sector público hacia los bancos y un rápido incremento en la mora.

Si bien hacia fines del año 2002 comenzaron a observarse indicios de una recuperación económica, que se iría consolidando en los años subsiguientes, la crisis dejó secuelas que pueden evidenciarse aún hoy.

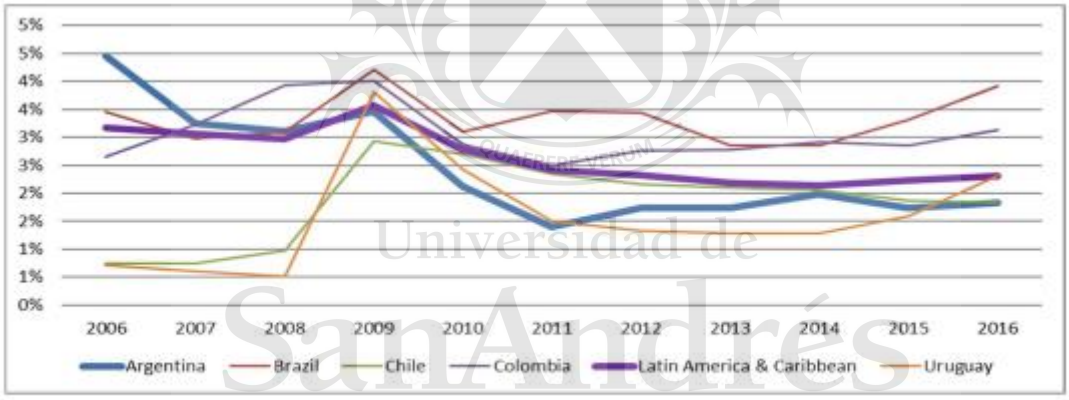
Durante el periodo postcrisis, el sistema financiero estuvo caracterizado por bajos niveles de apalancamiento, bajo otorgamiento de créditos y baja morosidad, manteniendo una magnitud muy pequeña respecto al PBI, sumado a una reducción en la cantidad de instituciones participantes en el sistema. La situación es peor si la comparamos a la de otros países de la región (Brasil, Chile, Colombia, Uruguay y la media de América Latina), donde se puede ver que Argentina ha registrado uno de los niveles de otorgamiento de créditos más bajos de la región. En los gráficos subsiguientes, puede observarse la evolución temporal de las variables otorgamiento de créditos y morosidad, en relación con los países enunciados previamente.

Gráfico 1 – Crédito interno al sector privado otorgado por bancos (% del PBI)[4]



Fuente: Elaboración propia en base a datos del banco mundial.

Gráfico 2 – Evolución NPL (Non-Performing Loans) respecto al total de créditos otorgados[5]



Fuente: Elaboración propia en base a datos del banco mundial.

Si bien el nivel inicial de morosidad de la cartera crediticia en Argentina denota un grado de deterioro sumamente elevado, cabe destacar su pronunciada mejoría como así también el hecho de que, a partir del año 2010, logra alcanzar uno de los niveles más bajos de la región. Un factor importante para esto es el énfasis que las autoridades pusieron en llevar a cabo los cambios normativos y regulatorios necesarios para lograr niveles de solvencia y liquidez que permitan

[4] Gráfico 1: Crédito interno al sector privado otorgado por bancos (% del PBI) - Periodo 1995 – 2016 – Base: Banco Mundial (Para el cálculo se incluyen todas aquellas instituciones capaces de tomar depósitos, sin tener en cuenta a la autoridad monetaria).

[5] Gráfico 2: Evolución NPL respecto al total de Créditos Otorgados – Periodo 2006 - 2016 – Base: Banco Mundial (La base contempla el total de créditos en situación de morosidad respecto a la cartera total de créditos otorgados)

encauzar nuevamente al sistema financiero en el sendero de la estabilidad, luego de la crisis del año 2001. Como resultado de las políticas adoptadas, se obtuvo un sistema financiero más sólido como así también más pequeño, en el cual tanto la oferta como la demanda se mostraron más conservadoras al momento de tomar un crédito. Vale destacar que no se ha vuelto a alcanzar el nivel de profundidad crediticia de los años anteriores a dicha crisis.

3.3 Variables a utilizar

En esta sección, enumeraré las variables a utilizar, agregando una breve descripción de su relación con la variable dependiente: Riesgo país. Esta será utilizada como variable "proxy" del estado de la economía para las estimaciones del filtro de Kalman. Entonces:

- **Riesgo país[6]:** Es un indicador que mide la diferencia de tasas que pagan los bonos del Tesoro de Estados Unidos (considerados libres de riesgo) contra las que pagan el resto de los países. La noción de riesgo país se vincula a la capacidad de pago de un Estado nación. Este índice mide el riesgo que existe en las operaciones internacionales que involucran a un determinado país, entendiendo como riesgo a la posibilidad de impago de este. Dado los factores que generan la estimación de este indicador, considero a esta variable como representativa del estado de la economía, a los fines de este trabajo.

- **EMAE[7]:** He seleccionado el EMAE (Estimador Mensual de Actividad Económica) para medir el nivel de actividad de la economía argentina. Se asume que existe una relación negativa entre el nivel de actividad y el riesgo país. Por ejemplo, frente a una caída del nivel de actividad se reducen los ingresos de las personas y compañías, ergo, cae el nivel de consumo e inversión. Eventualmente, el estado contará con menores

[6] Fuente: https://www.ambito.com/contenidos/riesgo-pais-historico.html
[7] Fuente: Indec: https://www.indec.gob.ar/nivel4_default.asp?id_tema_1=3&id_tema_2=9&id_tema_3=48

recursos para hacer frente al pago de su deuda ya que, frente a una caída del nivel de actividad, se recaudan menos impuestos. En el caso inverso, sube el nivel de actividad y el estado de la economía mejora, entonces existe un riesgo menor de impago de deuda por parte del estado.

- **IPC (Indice de Precios al Consumidor**)[8]: Los índices de precios al consumidor miden la variación de precios de los bienes y servicios representativos del consumo de los hogares residentes en la zona seleccionada en comparación con los precios vigentes en el año base. El aumento del nivel de precios tiene un efecto negativo en el repago de las deudas ya que, producto del aumento generalizado de los precios, se ve afectado el salario real de la población. Eventualmente, esto conlleva a un aumento de la morosidad de cartera.

- **Tasa de Desempleo[9]:** Un incremento de la tasa de desempleo deriva directamente en una caída del repago de las deudas, debido a que se ven afectados los ingresos de individuos y empresas. Consecuentemente, generaría un deterioro del estado de la economía.

- **Tipo de cambio real multilateral[10]:** Este índice mide el precio relativo de los bienes y servicios de una economía con respecto al de sus principales socios comerciales en función del flujo de comercio de manufacturas. Un tipo de cambio real multilateral muy atrasado, como bien tuvo la Argentina en la última década, implica una posterior degradación de la Balanza Comercial ya que se pierde nivel de competitividad. Esta pérdida de competitividad impacta en el nivel de actividad futura porque, eventualmente, se torna más difícil para los productores locales competir en el exterior. Es decir, se dificulta el acceso a los mercados internacionales por parte de los productores locales. Esta pérdida de mercados, sumada a la facilidad para productores extranjeros de competir

[8] Fuente: Indec
[9] Fuente: https://www.indec.gov.ar/series_historicas.asp?id_tema_1=4&id_tema_2=31&id_tema_3=58
[10] Fuente: http://www.bcra.gov.ar/PublicacionesEstadisticas/Indices_tipo_cambio_multilateral.asp

en el mercado local, conlleva una caída de las ventas/producción, deteriorando la capacidad de repago de los productores locales.

- **Nivel de Reservas:** Las Reservas Internacionales consisten en depósitos de moneda extranjera controlados por los bancos centrales y otras autoridades monetarias. Las reservas internacionales funcionan como indicador económico, mostrando los recursos de que dispone un país para hacer frente a sus obligaciones de pago en el extranjero. También, estos activos son usados por los bancos centrales para dar sustento a los pasivos emitidos en moneda extranjera.
  Un bajo nivel de reservas implica una baja capacidad de repago de la deuda externa. Es decir, un país que cuenta con bajos recursos para hacer frente a su propia deuda tiene más probabilidades de entrar en default.

- **Tasa BADLAR:** Se denomina así a la tasa de interés pagada a depósitos a plazo fijo, de 30 a 35 días, superiores a un millón de pesos. Es una tasa variable que calcula diariamente el Banco Central de la República Argentina (BCRA), en base a una muestra de tasas utilizadas en el ámbito de la ciudad de Buenos Aires (CABA) y el gran Buenos Aires (GBA). Se suele distinguir entre la Badlar Bancos Privados y Badlar Bancos Públicos. La misma es muy importante ya que marca el nivel de tasa de interés para el resto de las tasas de interés del sistema financiero argentino, incluidas tanto las que se pagan por los depósitos del público como las que se cobran por los créditos. Por otro lado, también se utiliza para como referencia para el pago de bonos (nacionales y provinciales) en pesos a tasa variable (Badlar más un porcentaje adicional). Este indicador tiene un impacto directo tanto en el repago de la cartera de créditos vigente, en el caso de créditos a tasa variable, como así también sobre la evolución del otorgamiento de créditos futuros. Por otro lado, el incremento de la tasa Badlar dificultaría el acceso a nueva financiación o tornaría más difícil el repago de un préstamo. Entonces, hay una relación positiva entre la tasa Badlar y la morosidad de cartera. También, frente a un aumento en las tasas de interés, eventualmente bajaría el nivel de

actividad porque el nivel de financiación al que uno podría acceder, ya sea por restricciones o por capacidad de repago de esta, sería inferior.

3.4 Metodología

Explicaré brevemente los pasos a seguir para la implementación de los métodos escogidos y descriptos en el capítulo 2.

El primer paso es la elección de las variables a utilizar junto con la justificación de dicha elección para estar seguros del camino que tomamos implementando estos métodos. Esto fue explicado en el inciso anterior.

El segundo paso es aplicar el filtro de Kalman. Algo necesario para esto es formular la matriz de transición de Kalman (T), a fin de conseguir la mejor estimación del vector que representa a la variable estado, *a* (tomaremos la siguiente fórmula como válida también para la mejor estimación del vector de medición):

$$\begin{bmatrix} a_{1t} \\ \vdots \\ a_{nt} \end{bmatrix} = \begin{bmatrix} T_{11} & \dots & T_{1n} \\ 1 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{1(t-1)} \\ \vdots \\ a_{n(t-1)} \end{bmatrix} + \begin{bmatrix} R_{11} & \dots & R_{1g} \\ \vdots & \ddots & \vdots \\ R_{n1} & \dots & R_{ng} \end{bmatrix} \cdot \begin{bmatrix} \eta_{1t} \\ \vdots \\ \eta_{gt} \end{bmatrix}$$

La diferencia entre ambas fórmulas de estimación, del vector estado y el vector medición, reside en reemplazar la matriz *T* por la matriz *Z* y el producto $R^*\eta$ por el vector ε.

Luego de esto, pueden suponerse las condiciones iniciales (media y varianza del estado) como dadas o no. En el caso de no conocer las condiciones iniciales, el algoritmo debe aplicarse con algunas modificaciones, explicadas en la bibliografía utilizada como referencia. Sin más, las recursiones del filtro de Kalman deberían hacer el resto del trabajo.

Vale mencionar la matriz de la ganancia de Kalman (K), ya que esta es la matriz clave que realiza el ajuste a la predicción al momento de recibir la nueva información. De acuerdo con las expresiones citadas en el capítulo anterior, la misma se escribe:

$$K_t = T_t P_t Z_t' F_t^{-1}$$

Esta fórmula tal vez no resulte muy intuitiva, por lo que agrego la siguiente versión simplificada, a modo de ejemplo, para una mejor comprensión por parte del lector:

$$K_t = \frac{P_t}{P_t + \sigma_\varepsilon^2} \qquad 0 \leq K \leq 1$$

Recordando que P_t representa la varianza del error de aproximación de α_t y que σ_ε^2 es la varianza del error de aproximación de y_t, podemos ver que la matriz de la ganancia de Kalman funciona como un ponderador de los errores de aproximación. Veamos las ecuaciones de actualización para entender cómo funciona:

$$a_{t+1} = Ta_t + K_t(y_t - a_t)$$
$$P_{t+1} = TP_t(T - K_tZ) + Q$$

En la primera ecuación actúa como un ponderador entre la variable estado y la observada. Cuanto más cercano está K de 0, mayor desconfianza tiene en la variable observada y más peso le da a la estimación del estado. Cuando K tiende a 1, se da el caso inverso. En la segunda ecuación, se ve de nuevo el caso de que al tender K a 1 se le da más confianza a la estimación de la variable estado y, por eso, se hace un ajuste negativo a la varianza de la misma.

Prosigamos con el tercer paso. Luego de obtener la variable estado, ya sea al momento *t* o una proyección futura de esta (dejaré la aclaración de esto para la siguiente subsección), es el momento de incluirla en el modelo de Vasicek:

$$\Pr(X_i < c \mid S) = N\left(\frac{N^{-1}(TTC\ PD) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

donde X_i es el valor aleatorio del que depende el default y está compuesto por el factor idiosincrático *Z* y el sistemático *S*, ambos relacionados por el factor de correlación entre activos ρ. La variable estado, calculada mediante el filtro de Kalman, es el factor sistemático S dentro del modelo de Vasicek. El ajuste se realiza sobre la probabilidad de default *constante a través del ciclo* (TTC PD), transformándola en una variable condicionada al momento del ciclo de crédito.

Este ajuste es necesario realizarlo a cada uno de los valores de la matriz de transición que representen probabilidades de default. En general, estos son los

valores de transición desde la calificación que analicemos hacia la peor categoría de la grilla. Se consideran como probabilidad de default ya que no se espera un repago futuro de las obligaciones al caer en dicha categoría.

Luego de realizar estos ajustes, uno esta listo para proseguir con el resto del proceso hacia la obtención de la pérdida esperada (ECL). El resto de este proceso no forma parte del enfoque del trabajo.

Proyección utilizando el Filtro de Kalman

Durbin y Koopman explican que el proceso utilizado para las proyecciones es considerar las mismas como observaciones "faltantes" pero posteriores a las últimas observadas, y utilizar las técnicas para tratar con observaciones faltantes. Se continúa iterando, pero sin incluir nuevas observaciones. Es decir, se siguen realizando las estimaciones únicamente a partir de los valores pasados de la variable de estado no observable. Los resultados siguen siendo válidos para ser MVLUE (estimador lineal insesgado de varianza mínima, por sus siglas en inglés) y para los supuestos Bayesianos. Este tratamiento simple de los datos faltantes es uno de los grandes atractivos de los métodos de espacio de estado utilizados para análisis de series de tiempo.
La explicación de esta metodología queda fuera del enfoque de nuestro trabajo, pero para más detalles se puede recurrir al libro de Durbin y Koopman.

Observaciones
En cuanto a las observaciones, dado que utilizo el filtro de Kalman para estimar variables normales estandarizadas, considero más adecuado estandarizar primero las variables utilizadas que estandarizar luego las estimaciones obtenidas con el filtro. Dada la conformidad con los resultados, puede decirse que la decisión no es incorrecta. Por otro lado, si decidiera utilizar los datos para realizar una regresión lineal, el hecho de tener las variables estandarizadas permitiría evaluar no sólo la significancia estadística sino también la económica. Además, puede ser que para determinadas variables la media de esta no sea el valor de referencia más significativo. Centrándome específicamente en las mediciones de la variable "Riesgo País" (considerada como proxy de la variable

estado del Ciclo del Crédito), tal vez sería más prudente tomar la media de Latinoamérica y no la de Argentina únicamente. Encontraríamos un benchmark de economías y riesgos similares, pero que brindaría un valor numérico con un contexto más amplio al momento de considerarlo como determinante para decidir cuáles valores vamos a pensar como positivos o negativos dentro del índice del Ciclo del Crédito.

3.5 Resultados

Para trabajar con los modelos y fórmulas explicados anteriormente, utilicé la matriz de transición e incumplimiento provista por Fix Scr – Fitch Ratings:

"FIX – Finanzas Corporativas – Matrices de transición de Calificaciones Nacionales", con los valores promedio a 1 año de los años 2002-2016. El documento completo puede encontrarse en:

http://www.fixscr.com/uploads/1486756307589e19d32720c.pdf.

Veamos dicha matriz:

Rating	AAA	AA	A	BBB	BB	B	CCC	CC	C	D	R
AAA	0.79	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08
AA	0.03	0.86	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.05
A	0.00	0.09	0.76	0.05	0.00	0.00	0.00	0.00	0.00	0.02	0.08
BBB	0.00	0.01	0.12	0.69	0.02	0.01	0.00	0.00	0.01	0.02	0.12
BB	0.00	0.00	0.06	0.12	0.50	0.03	0.00	0.00	0.00	0.03	0.27
B	0.00	0.00	0.00	0.06	0.19	0.50	0.13	0.00	0.00	0.00	0.13
CCC	0.00	0.00	0.00	0.00	0.00	0.17	0.50	0.00	0.00	0.17	0.17
CC	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.20	0.00	0.60	0.00
C	0.00	0.00	0.00	0.00	0.20	0.20	0.00	0.00	0.20	0.20	0.20

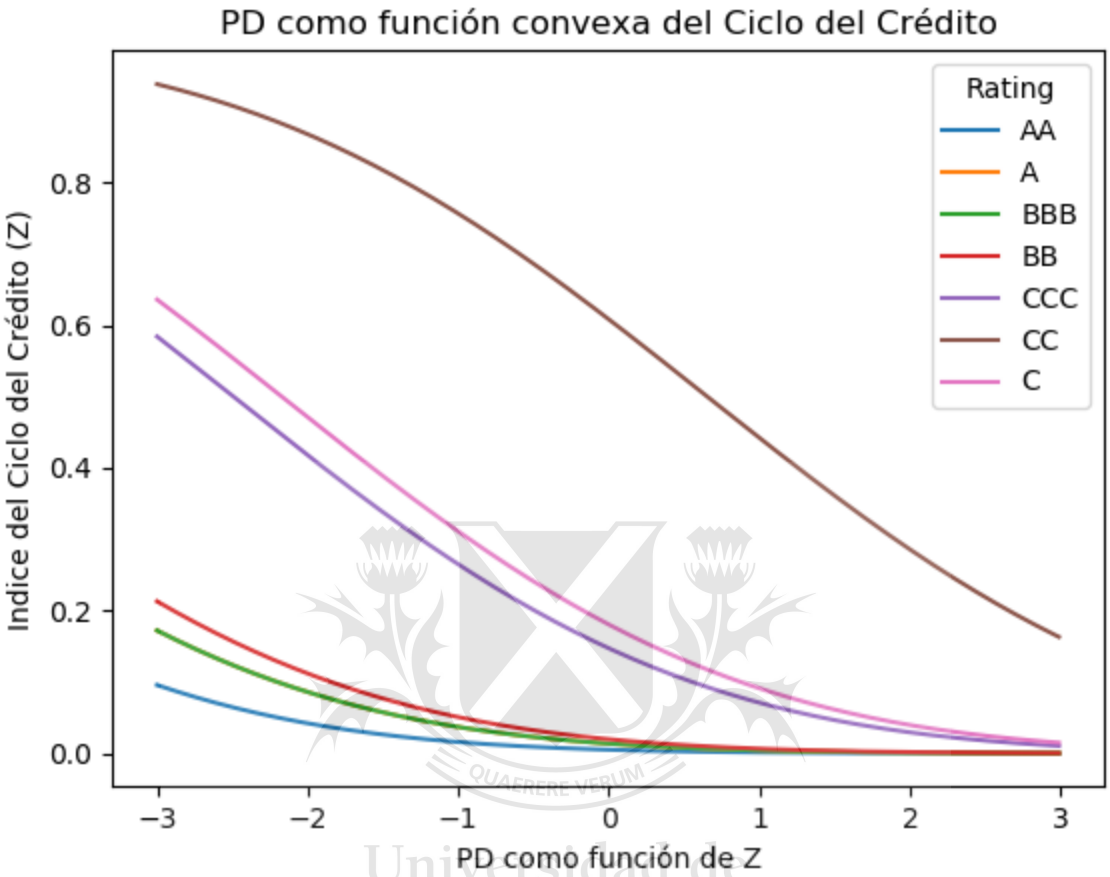
En esta matriz vemos los valores cruzados entre los diferentes ratings corporativos, significando estos números la probabilidad de cambiar de la categoría que nos indica la fila a la categoría indicada por la columna. La categoría "D" representa a las empresas que entraron en default o incumplimiento, mientras que la categoría "R" representa el porcentaje de empresas de cierta categoría que se retiraron del mercado y no deben ser consideradas para la tabla del año siguiente. Por esto, ninguna de estas dos categorías figura en la columna de "Ratings" ya que no pueden cambiar de calificación.

En la tabla subsiguiente se observan los valores de la probabilidad de default "estática a través del ciclo" de la matriz de transición utilizada (únicamente para las categorías cuya PD es distinta de 0) y, el ajuste por convexidad realizado para cada fecha considerada respecto del valor del índice del ciclo de crédito estimado a partir del Filtro de Kalman.

Fecha (Momento "t")	Indice del Ciclo del Crédito	PD "AA" Ajustada	PD "A" Ajustada	PD "BBB" Ajustada	PD "BB" Ajustada	PD "CCC" Ajustada	PD "CC" Ajustada	PD "C" Ajustada
		Base 0.90%	Base 2.10%	Base 2.10%	Base 2.90%	Base 16.70%	Base 60.00%	Base 20.00%
01/04/2008	-0.53	0.96%	2.37%	2.37%	3.34%	20.48%	69.08%	24.52%
01/07/2008	-0.52	0.95%	2.35%	2.35%	3.31%	20.35%	68.92%	24.38%
01/10/2008	-0.56	0.98%	2.43%	2.43%	3.42%	20.76%	69.43%	24.84%
01/01/2009	-0.44	0.87%	2.17%	2.17%	3.07%	19.44%	67.75%	23.37%
01/04/2009	-0.36	0.79%	2.00%	2.00%	2.84%	18.52%	66.52%	22.34%
01/07/2009	-0.31	0.74%	1.90%	1.90%	2.71%	17.94%	65.73%	21.70%
01/10/2009	-0.14	0.61%	1.59%	1.59%	2.28%	16.11%	63.03%	19.63%
01/01/2010	-0.01	0.52%	1.39%	1.39%	2.01%	14.84%	61.01%	18.19%
01/04/2010	0.37	0.32%	0.91%	0.91%	1.35%	11.42%	54.71%	14.25%
01/07/2010	0.64	0.23%	0.66%	0.66%	1.00%	9.37%	50.17%	11.83%
01/10/2010	0.73	0.20%	0.60%	0.60%	0.90%	8.76%	48.67%	11.11%
01/01/2011	0.26	0.37%	1.03%	1.03%	1.52%	12.38%	56.61%	15.36%
01/04/2011	0.03	0.50%	1.33%	1.33%	1.94%	14.48%	60.41%	17.78%
01/07/2011	0.21	0.40%	1.10%	1.10%	1.61%	12.84%	57.48%	15.89%
01/10/2011	0.10	0.46%	1.23%	1.23%	1.79%	13.79%	59.21%	16.99%
01/01/2012	0.02	0.50%	1.34%	1.34%	1.95%	14.55%	60.51%	17.85%
01/04/2012	-0.04	0.54%	1.43%	1.43%	2.07%	15.10%	61.44%	18.49%
01/07/2012	-0.48	0.91%	2.26%	2.26%	3.19%	19.89%	68.34%	23.87%
01/10/2012	-0.69	1.14%	2.77%	2.77%	3.86%	22.41%	71.37%	26.65%
01/01/2013	-0.78	1.27%	3.03%	3.03%	4.21%	23.62%	72.71%	27.97%
01/04/2013	-0.50	0.92%	2.30%	2.30%	3.25%	20.12%	68.62%	24.12%
01/07/2013	-0.31	0.74%	1.90%	1.90%	2.70%	17.93%	65.72%	21.68%
01/10/2013	-0.45	0.88%	2.19%	2.19%	3.11%	19.57%	67.92%	23.51%
01/01/2014	-0.40	0.82%	2.07%	2.07%	2.94%	18.91%	67.05%	22.77%
01/04/2014	0.03	0.50%	1.33%	1.33%	1.94%	14.49%	60.42%	17.79%
01/07/2014	0.79	0.19%	0.56%	0.56%	0.84%	8.36%	47.67%	10.64%
01/10/2014	1.92	0.04%	0.13%	0.13%	0.21%	3.18%	29.70%	4.27%
01/01/2015	1.62	0.06%	0.20%	0.20%	0.31%	4.21%	34.31%	5.57%
01/04/2015	0.47	0.29%	0.81%	0.81%	1.21%	10.67%	53.12%	13.36%
01/07/2015	-0.25	0.69%	1.78%	1.78%	2.54%	17.25%	64.73%	20.91%
01/10/2015	-0.37	0.80%	2.02%	2.02%	2.86%	18.60%	66.64%	22.43%

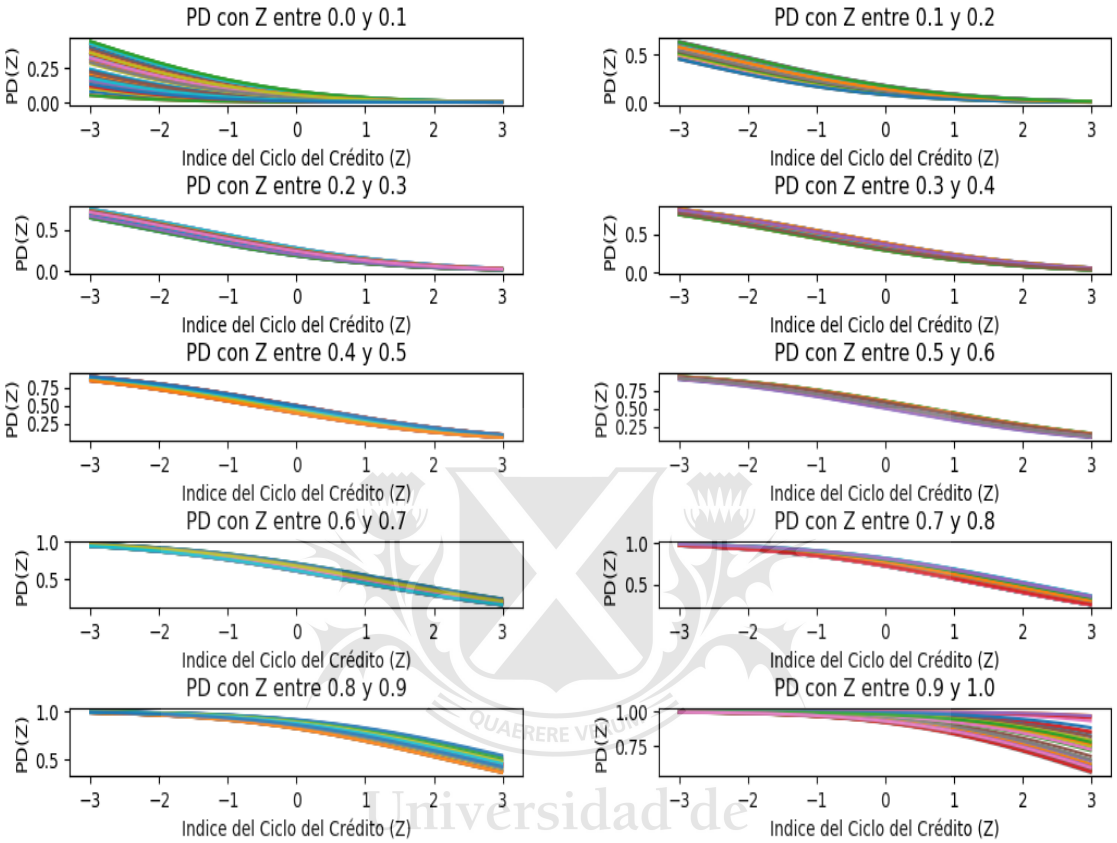
Esta tabla nos brinda la serie temporal de la probabilidad de default ajustada en las fechas consideradas para el trabajo. Pero, dejando de lado el factor temporal, creé un gráfico con los ajustes posibles que pueden impactar en una PD utilizando la fórmula de Vacisek y considerando al factor sistemático como una N(0, 1) acotada entre los valores -3 y 3, con pasos de 0.01. Esto último se debe

a que los valores de una N(0, 1) que excede los 3 desvíos estándares respecto de la media no tienen significancia estadística.



En este gráfico observamos claramente la convexidad que genera el Ciclo del Crédito en la distribución de la PD. Está claro que las variaciones negativas de dicho ciclo producen un impacto negativo mayor en dicha probabilidad que los impactos positivos generados por los buenos momentos del ciclo. Curiosamente, en el caso de las empresas con rating “CC”, esto último no puede notarse tan claramente. Esto despertó mi interés, por lo que decidí realizar simulaciones de probabilidades de default para tratar de analizar la mayor cantidad posible de distribuciones de PD “base” ajustadas con el modelo de Vasicek. Nuevamente utilicé los valores de una N(0, 1) entre -3 y 3 con pasos de 0.01 de magnitud para realizar los ajustes. Por último, luego de calcular los ajustes para cada una de las probabilidades simuladas, decidí agrupar los gráficos de las distribuciones en 10 categorías distintas. Cada categoría representa un intervalo de 0.1 de amplitud en la probabilidad de default, siendo el primero entre 0 y 0.1 y el último

entre 0.9 y 1. De este modo, podemos ver la incidencia de la convexidad para los diferentes intervalos de probabilidad de default. Veamos los diez gráficos a continuación:



Aquí se nota como el efecto de la convexidad se va invirtiendo a medida que la probabilidad de default “base” es mayor. Esto se debe a que el impacto se vuelve proporcionalmente menor a medida que la PD TTC crece y, además, cuando esta se acerca a 1 el efecto negativo tiene un tope dado que el valor máximo de una probabilidad es 1. Esto último puede verse claramente en los últimos dos gráficos donde parece que sólo hay un efecto positivo del ciclo del crédito sobre la PD TTC.

3.6 Conclusiones

El trabajo realizado ha sido muy enriquecedor y ha tenido resultados positivos en cuanto al aprendizaje y el análisis de los modelos. A simple vista, sobre todo de los gráficos, pareciera ser que los números responden directamente a los planteos teóricos que los sustentan. Pero parte de lo más valioso se esconde en

los detalles, que tal vez sean contemplaciones que los autores de los modelos no han tenido, no han notado o han pensado que el beneficio era más grande que reparar en esos detalles.

Observando el gráfico de distribuciones de las PD de la matriz de transición utilizada, sin tener en cuenta la tabla de valores, vemos que el impacto negativo es mayor claramente en casi todos los casos (este otro caso ha sido explicado en el punto anterior). A simple vista todo se encuentra funcionando correctamente. Pero, adentrándome en la tabla he notado que, para los impactos negativos pequeños[11], la probabilidad de default es menor que la original. Esto se debe a que el efecto negativo, hasta un cierto umbral, no compensa el efecto de las "correlaciones" en el divisor de la fórmula de Vasicek. Intuitivamente esto parece un error, pero recordemos que la nueva probabilidad de default se calcula aislando el único factor que desconocido: la probabilidad de default idiosincrática. Es decir, antes teníamos una probabilidad de default determinada que provenía de la incertidumbre que rodeaba a ambos factores, el sistemático y el idiosincrático. Dada la nueva información sobre el factor sistemático, dicha incertidumbre se reduce y la probabilidad de default debe ser recalculada como una probabilidad condicionada a un determinado valor de dicho factor. En cuanto a los valores numéricos, no quiere decir que el umbral del impacto negativo sea siempre el mismo ya que este depende de la magnitud de la correlación entre los activos considerada. Cuanto más baja sea esa correlación, más pequeño será el umbral, ya que el factor sistemático tendría una menor ponderación y no habría un reescalamiento tan grande porque la PD comenzaría a depender casi exclusivamente del factor idiosincrático. Para concluir, debemos considerar como nueva PD "base" cuando realizamos el ajuste a la que obtenemos con un S = 0 y, a partir de ese valor, evaluar el impacto de los movimientos positivos y negativos del ciclo.

Por otro lado, también cabe mencionar el hallazgo logrado gracias a las simulaciones. Ya hemos mencionado y explicado brevemente el por qué de la inversión de la convexidad a medida que la PD TTC se incrementa. También

[11] La magnitud del valor del índice depende del valor de correlación que utilicemos. En este caso, con una correlación de 0.15, los valores del índice entre -0.5 aproximadamente y 0 también tienen una menor probabilidad de default.

cabe destacar que con una PD entre 0.5 y 0.7 aproximadamente, pareciera no haber un efecto significativo de convexidad. Entonces, evaluemos dos casos:

- PD TTC se encuentra entre 0.5 y 0.7,
- PD TTC es mayor a 0.7.

En el primer caso, habría que analizar si realmente vale la pena aplicar el modelo o es simplemente un desperdicio de recursos y, tal vez, este intervalo merezca otra clase de análisis. Es cierto que, por lo explicado al principio de esta sección, hay un corrimiento en lo que llamamos la “nueva” PD base. Pero si no hay convexidad significativa, probablemente, podría estimarse un corrimiento fijo de las PD TTC en lugar de utilizar modelos más complejos. Por otro lado, para las PD que entran en el segundo caso hay que tener un cuidado especial. Si uno aplica este modelo no debería tener problemas. Pero si uno aborda los efectos de la parte negativa del ciclo del crédito para estas PD de otra forma, podría terminar sobreestimando el impacto negativo o subestimando el impacto positivo. Por ejemplo, en los análisis de escenarios suelen plantearse resultados más negativos (y positivos) que el escenario base y se estiman probabilidades de default más pesimistas (u optimistas). El problema es que, bajo el modelo de Vasicek, la probabilidad de default viene dada por la incapacidad de repagar las deudas con los activos que poseemos y la evolución de sus valores en el tiempo. Por esto, afectar negativamente los flujos a nuestro favor y además considerar una PD más grande podría, bajo el modelo de Vasicek, derivar en una PD mucho mayor de forma injustificada. Aquí no solo sobreestimaríamos la PD sino todo lo que la incluye como variable para sus cálculos, ya sean reservas, primas a cobrar, etc. Además, a partir de ciertos niveles de la PD no tendría sentido subdividir los escenarios pesimistas ya que no podrían generar un impacto más negativo que el existente. Por otro lado, por conservadurismo suelen considerarse como menores los impactos positivos respecto de los negativos. No se suele justificar esto con la convexidad, si bien esta propiedad lo demuestra para gran parte de los casos. Pero, para niveles altos de PD, deberían plantearse escenarios positivos sin compararlos con los escenarios negativos. Los límites en el valor de la PD, que invierten la convexidad cuando esta es muy alta, implican que pierda sentido tomar como referencia el valor absoluto del impacto

negativo para elaborar la estimación de un escenario positivo. Creo yo que debería realizarse un análisis independiente para estos casos.

Como conclusión general, es importante entender bien los modelos aplicados tanto teórica como prácticamente. La teoría es importante pero los números (ya sean por definiciones matemáticas o por lógica) deben ser analizados. Subdividir en intervalos el análisis de las PD TTC parece una buena práctica, aunque la granularidad de dicho análisis deberá determinarla cada entidad en particular, no solo en base a recursos disponibles sino para que la calidad (y cantidad) de la información no perjudique el análisis.

CAPITULO 4
BIBLIOGRAFIA

- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Engineering, Transactions ASMA, Series D 82*, 35–45.
- Koopman, S. J. and J. Durbin (2012). *Time Series Analysis by State Space Methods 2nd Ed.*
- Schwaab, B. (2011). *Credit Risk and State Space Methods, Academish Proefschrift.*
- Chawla, G. and Forest Jr., L. R. (2016). *Convexity and Correlation Effects in Expected Credit Loss calculations for IFRS9/CECL and Stress Testing, vol. 9 Number 4, Autumn 2016 edition of Journal of Risk Management in Financial Institutions.*
- Chatterjee, S. (2015). *Modelling credit risk*, *Center for Central Banking Studies, Bank of England.*
- Van Biezen, M. *Special Topics – The Kalman Filter (canal de youtube: https://www.youtube.com/channel/UCiGxYawhEp4QyFcX0R60YdQ)*

ANEXO A

MODELO DE VALUACIÓN DE PORTAFOLIO DE PRÉSTAMOS: VASICEK

Vasicek aplicó al valor de los activos de las empresas lo que se conoce como el movimiento geométrico Browniano. Expresa la variación del valor de los activos de la empresa *i* con la siguiente ecuación diferencial estocástica,

$$dA_i = \mu_i A_i dt + \sigma_i A_i dx_i$$

Donde A_i es el valor de los activos de la empresa *i*, μ_i y σ_i son el drift y la volatilidad de aquel valor, y x_i es un proceso de Wiener o movimiento Browniano, es decir, un movimiento aleatorio continuo en el tiempo en el cual el cambio sobre cualquier periodo finito de tiempo esta distribuido normalmente con media cero y varianza igual a la longitud de dicho periodo, y las variaciones del mismo entre periodos de tiempo disjuntos son independientes entre sí. Resolviendo esta ecuación diferencial estocástica, uno obtiene el valor de los activos de la firma *i* en el momento T:

(1) $A_i(T) = e^{A(0)+\mu_i T-\frac{1}{2}\sigma_i^2 T+\sigma_i\sqrt{T}X_i}$

Suponemos que B es el monto de las obligaciones de la empresa, por lo que la misma defaultea si el valor de los activos en el momento final no logra cubrirlas, A(T) < B. Entonces, la probabilidad de dicho evento es:

(2) $P[A_i(T) < B_i] = P[X_i < c_i] = N(c_i) = p^*$

Suponiendo que B es el nivel mínimo que debería tomar A(T) en la ecuación (1) para no defaultear y, sabiendo que la única variable que no es determinística en su estimación es *X*, llamamos *c* al valor mínimo que debe tomar *X* para que la empresa no caiga en default. Resolviendo $A_i(T) < B_i$, todos los demás términos pueden simplificarse y la probabilidad de default dependerá exclusivamente del valor que tome *c.* El valor de *c* se determina fácilmente reemplazando A(T) por B en la ecuación (1) y resolviendo.

$$c_i = (\ln(B) - A(0) - \mu_i T + \frac{1}{2}\sigma_i^2 T) \cdot \frac{1}{\sigma_i\sqrt{T}}$$

La correlación entre defaults se introduce asumiendo correlación en los procesos de A_i y, en consecuencia, en los valores finales $A_i(T)$. En particular, se asume que los X_is de la ecuación (1) están correlacionados de a pares de acuerdo con el factor ρ.

Siendo normales y equi-correlacionadas, cada variable aleatoria X_i puede ser representada como la suma de otras dos variables normales: una común para todas las firmas y la otra idiosincrática:

$$X_{it} = S_t\sqrt{\rho} + Z_{it}\sqrt{1-\rho}$$

Con $S \sim N(0,1)$, $Z_i \sim N(0,1)$. Entonces, la probabilidad de default de la empresa *i* puede ser escrita también de la siguiente forma

$$P[A_i(T) < B_i] = P[X_i < c_i]$$
$$= P[X_i < N^{-1}(p^*)]$$
$$= P\left[\left(S\sqrt{\rho} + Z_i\sqrt{1-\rho}\right) < N^{-1}(p^*)\right]$$
$$(3) = P\left[Z_i < \frac{c_i - S\sqrt{\rho}}{\sqrt{1-\rho}}\right] = N\left(\frac{N^{-1}(p^*) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right)$$

p* en la ecuación (2) es la pérdida promedio a través del ciclo (TTC). Vasicek parte de suponer una cartera de todos préstamos iguales y sin recupero, por lo que el promedio de la pérdida sería equivalente a la probabilidad de default. Es decir, cuando el número de préstamos tiende a infinito, la proporción de las pérdidas tiende a su media, la PD ajustada por S. Esta última está calculada por la ecuación (3).

La siguiente ecuación muestra que la proporción de los préstamos que defaultean tienen la siguiente función de distribución de probabilidad:

$$P[p(S) \leq x] = P[S \geq p^{-1}(x)]$$
$$(4) = N\left(-p^{-1}(x)\right) = N\left(\frac{\sqrt{1-\rho} \cdot N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right)$$

La ecuación que precede a la (4) puede no ser muy intuitiva por lo que cabe hacer un par de aclaraciones:

- p(S): Es la probabilidad de default ajustada por el estado *S;*
- *S* es el valor del estado *S* y no es la inversa de p(S)

- Siendo la normal estándar una función simétrica, $P[S \geq p^{-1}(x)]$ es igual a $P[S \leq -p^{-1}(x)] = N\left(-p^{-1}(x)\right)$.

Para una explicación más detallada del modelo de Vasicek pueden leerse los documentos referidos en la bibliografía.

ANEXO B

KALMAN FILTER: ESTIMACIÓN DE PARÁMETROS

En este anexo describiré el caso en el cual las condiciones iniciales son conocidas: a_0 y P_0. Para ver el caso en el cual estas condiciones son difusas puede recurrirse a Durbin y Koopman (2012).

Muchas veces se asumen como conocidas las matrices del sistema Z_t, H_t, R_t, T_t y Q_t para $t = 1, 2, \dots, T$. Pero hay una gran cantidad de algoritmos numéricos disponibles para maximizar la verosimilitud con respecto a parámetros desconocidos. Algunos de estos están basados en el método de Newton que resuelve la ecuación.

(1) $$\partial_1(\psi) = \frac{\partial logL(Y_n|\psi)}{\partial\psi} = 0,$$

usando las series de Taylor de primer orden

(2) $$\partial_1(\psi) \cong \partial_1^*(\psi) + \partial_2^*(\psi)(\psi - \psi^*)$$

para un cierto valor ψ^*, donde

$$\partial_1^*(\psi) = \partial_1(\psi)|_{\psi=\psi^*} \qquad \partial_2^*(\psi) = \partial_2(\psi)|_{\psi=\psi^*}$$

Siendo

(3) $$\partial_2(\psi) = \frac{\partial^2 logL(Y_n|\psi)}{\partial\psi\partial\psi'}$$

Igualando (2) a cero obtenemos un nuevo valor estimado ψ^{**} de la expresión

$$\psi^{**} = \psi^* - \partial_2^*(\psi)^{-1}\partial_1^*(\psi).$$

Este proceso se repite hasta la convergencia. Cabe mencionar que debe recalcularse el valor de cada parámetro del vector antes de iniciar la siguiente iteración con los nuevos valores. El *gradiente* $\partial_1(\psi)$ marca la dirección del paso tomado hacia el óptimo y la matriz Hessiana $\partial_2(\psi)$ modifica el tamaño de dicho paso. Es posible excederse del punto óptimo yendo en la dirección determinada por el vector

$$\pi^*(\psi) = -\partial_2^*(\psi)^{-1}\partial_1^*(\psi),$$

por lo que es práctica común incluir un factor escalar de velocidad de convergencia en el proceso de optimización. Obteniendo el algoritmo

$$\psi^{**} = \psi^{*} + s\pi^{*}(\psi),$$

encontrándose el valor ideal de s, generalmente, entre 0 y 1. De este modo, vemos que este método es similar al método del gradiente descendiente.

ANEXO C

KALMAN FILTER: EJEMPLO PRACTICO

Veamos un ejemplo de Kalman Filter para estimar la ubicación de un objeto. Sea *"y"* el vector de variables observadas, "*x*" el de variables estado, "*u*" el de variables control y suponiendo $y_0 = x_0$:

<u>Supuestos del ejemplo:</u>

Medidas de ubicación: $y_{10} = 4000$, $y_{11} = 4260$

Medidas de velocidad: $y_{20} = 280$, $y_{21} = 282$

Variable control: $u = 2$

Paso temporal: $\Delta t = 1$

Errores:

De estimación: $\Delta P_{x_1} = 20$; $\Delta P_{x_2} = 5$

De observación: $\Delta y_1 = 25$; $\Delta y_2 = 6$

<u>Proceso de estimación</u>

Completaremos las ecuaciones de transición y de actualización, además de las ecuaciones intermedias para lo mismo. El subíndice de las variables se refiere al momento en el tiempo en que se encuentran. En el caso de las variables con subíndice "1;0", significa que es la proyección que se estima para el momento "1" con la información disponible hasta el momento "0". Es decir, previo a que se obtengan las medidas de las variables en el momento "1".

Paso de estimación: *Se realizan las estimaciones para el momento "1" en base a la información con la que se cuenta en el momento "0".*

Ecuación de transición proyectada: $x_{1;0} = Zx_0 + Bu_1 + w_k$

$$x_{1;0} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} \cdot 2 + 0 = \begin{bmatrix} x_{10} + x_{20}\Delta t + \Delta t^2 \\ x_{20} + 2\Delta t \end{bmatrix} = \begin{bmatrix} 4000 + 280 + 1 \\ 280 + 2 \end{bmatrix} = \begin{bmatrix} 4281 \\ 282 \end{bmatrix}$$

Proyección de la varianza de la estimación:

$$P_{1;0} = ZP_0Z^T + Q_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 400 & 0 \\ 0 & 25 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + 0 = \begin{bmatrix} 425 & 25 \\ 25 & 25 \end{bmatrix} \cong \begin{bmatrix} 425 & 0 \\ 0 & 25 \end{bmatrix}$$

Ganancia de Kalman (factor de ponderación en base a los errores de las estimaciones de estados y de la medición de las variables):

$$K = \frac{P_{1;0}Z^T}{ZP_{1;0}Z^T + H}, \qquad Z = I \Rightarrow K = \frac{P_{1;0}}{P_{1;0} + H} = \frac{\begin{bmatrix} 425 & 0 \\ 0 & 25 \end{bmatrix}}{\begin{bmatrix} 425 & 0 \\ 0 & 25 \end{bmatrix} + \begin{bmatrix} 625 & 0 \\ 0 & 36 \end{bmatrix}} = \begin{bmatrix} 0{,}405 & 0 \\ 0 & 0{,}410 \end{bmatrix}$$

Paso de actualización: *Se ajustan las estimaciones realizadas una vez que observamos los valores de las variables en el momento "1".*

$$y_1 = \begin{bmatrix} 4260 \\ 282 \end{bmatrix}$$

$$x_1 = x_{1;0} + K[y_1 - Zx_{1;0}] = \begin{bmatrix} 4281 \\ 282 \end{bmatrix} + \begin{bmatrix} 0{,}405 & 0 \\ 0 & 0{,}410 \end{bmatrix} \left\{ \begin{bmatrix} 4260 \\ 282 \end{bmatrix} - I \cdot \begin{bmatrix} 4281 \\ 282 \end{bmatrix} \right\} = \begin{bmatrix} 4272{,}5 \\ 282 \end{bmatrix}$$

$$P_1 = (I - KZ)P_{1;0} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0{,}405 & 0 \\ 0 & 0{,}410 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} 425 & 0 \\ 0 & 25 \end{bmatrix} = \begin{bmatrix} 253 & 0 \\ 0 & 14{,}8 \end{bmatrix}$$

A partir de aquí podemos ver cómo funciona el ciclo recursivo. Para el próximo paso x_1 y P_1 serán los "pasos previos" para generar la estimación del periodo siguiente, que luego se complementará con el siguiente valor observado, y_2, etc.

ANEXO D

TABLA DE DATOS

Cuatrimestre	Riesgo País	Tasa de desmpleo	BADLAR	EMAE	Nivel de Reservas (mill. USD)	IPC	TCRM
3Q2015	589.00	5.90	23.07	147.05	32471.00	1319.59	75.12
2Q2015	624.00	6.60	22.08	151.84	33847.00	1243.63	83.50
1Q2015	628.00	7.10	22.10	157.48	31505.00	1177.31	84.32
4Q2014	726.00	6.90	21.92	134.00	31337.00	1108.54	91.97
3Q2014	753.00	7.50	21.07	143.35	27921.00	1040.38	98.78
2Q2014	697.00	7.50	21.72	145.42	29339.00	969.38	108.43
1Q2014	802.00	7.10	25.32	152.34	26989.00	899.00	112.39
4Q2013	800.00	6.40	20.69	137.97	30586.00	820.77	102.56
3Q2013	1039.00	6.80	18.21	148.69	34825.00	738.38	100.09
2Q2013	1189.00	7.20	16.71	150.73	37132.00	691.83	97.83
1Q2013	1321.00	7.90	14.89	155.94	40413.00	639.74	101.69
4Q2012	949.00	6.90	14.96	136.52	43232.00	617.95	102.16
3Q2012	880.00	7.60	14.57	147.23	45089.00	583.34	101.75
2Q2012	1071.00	7.20	13.10	146.86	46215.00	553.64	100.99
1Q2012	902.00	7.10	12.35	145.52	47241.00	529.37	108.21
4Q2011	925.00	6.70	15.48	136.19	46231.00	497.63	109.05
3Q2011	1018.00	7.20	15.52	145.05	48409.00	473.76	111.03
2Q2011	556.00	7.30	11.00	146.38	51749.00	448.17	126.26
1Q2011	513.00	7.40	10.64	151.04	51302.00	428.20	127.44
4Q2010	483.00	7.30	10.44	133.08	52179.00	405.96	125.44
3Q2010	662.00	7.50	10.15	135.14	51236.00	381.10	131.08
2Q2010	846.00	7.90	9.92	139.51	49325.00	361.38	127.06
1Q2010	646.00	8.30	9.38	145.77	47472.00	344.13	132.92
4Q2009	631.00	8.40	9.83	121.51	48068.00	320.45	143.53
3Q2009	797.00	9.10	11.81	127.08	45351.00	303.35	150.69
2Q2009	1040.00	8.80	13.07	127.36	46035.00	294.15	149.83
1Q2009	1893.00	8.40	12.91	127.99	46390.00	285.62	137.69
4Q2008	1681.00	7.30	15.98	116.76	46408.00	275.93	132.51
3Q2008	1005.00	7.80	15.53	130.40	47152.00	263.85	133.98
2Q2008	629.00	8.00	14.35	136.64	47524.00	254.31	146.51
1Q2008	561.00	8.40	8.93	145.28	50426.00	243.96	156.01

ANEXO E

CÓDIGO DE PYTHON PARA LA EJECUCIÓN DEL MODELO

```
# -*- coding: utf-8 -*-
"""
Ejecución de Kalman Filter
El propósito de este código es estimar la variable estado del ciclo de crédito para incluirla en el modelo
de cálculo de probabilidad de default de Vasicek. Luego se realizan gráficos para analizar los resultados
obtenidos y hacer comparaciones con las probabilidades de default originales.

Origen de los Datos utilizados: Provenientes de Argentina

Variables utilizadas:
   - rp: Riesgo país (Variable proxy de la variable estado)
   - emae: Estimador mensual de actividad económica
   - badlar: Tasa Badlar mensual
   - tdd: Tasa de desempleo
   - nr: Nivel de reservas
   - ipc: Indice de precios al consumidor
   - tcrm: Tipo de cambio real multilateral

Created on Sun Jan 13 20:54:13 2019

@author: Martin Senuy
"""
# Aquí se cargan los paquetes y funciones necesarias para el trabajo. La descarga e instalación de los
# mismos en la computadora corre por cuenta del usuario.
from pykalman import KalmanFilter
import numpy as np
import pandas as pd
from scipy.stats import norm
import math
import matplotlib.pyplot as plt

# ==============================================================================
# Importar información: Mediciones observables para aplicar Kalman Filter
# ==============================================================================
folder = r'C:\Users\Martin\Desktop\TESIS\Bases de datos'
filename = '\BaseDeDatosConjunta.xlsx'
path = folder + filename
Dataset = pd.read_excel(path)

# ==============================================================================
# Creación de la matriz de Transición y Default
# ==============================================================================
# Matriz de transición y default utilizada: Fix Scr, Fitch Ratings
# http://www.fixscr.com/uploads/1486756307589e19d32720c.pdf
```

```
transition_matrix = np.array([[100, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], [20, 75, 2.5, 0, 0, 0, 0, 0, 0, 0, 2.5],
                    [0, 8.5, 85.1, 6.4, 0,  0,  0,  0,  0,  0, 0], [0, 0, 12.5, 75, 6.3, 3.1, 0, 0, 0, 0, 3.1],
                    [0, 0, 0, 0, 100, 0, 0, 0, 0, 0, 0], [0, 0, 0, 0, 0, 100, 0, 0, 0, 0, 0],
                    [0, 0, 0, 0, 0, 100, 0, 0, 0, 0, 0], [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
                    [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]])

ratings = ['AAA', 'AA', 'A', 'BBB', 'BB', 'B', 'CCC', 'CC', 'C', 'D', 'R']

transition_matrix = pd.DataFrame(transition_matrix, columns = ratings, index = ratings[0:-2])

# ==============================================================================
# Ejecución del filtro de Kalman
# ==============================================================================
varNums = Dataset.reset_index().iloc[:-1, 2:] # Variables numericas para KF
vN_mean = varNums.apply(np.mean)
vn_sd = varNums.apply(np.std)
vn_Standardised = ((varNums - vN_mean)/vn_sd)
# Se estandarizan las variables para poder aplicar Kalman Filter más adecuadamente
varNums = vn_Standardised.values.tolist()

# Aplicación del filtro de Kalman
kf = KalmanFilter(n_dim_obs = 7)
kf = kf.em(varNums, n_iter = 1).smooth(varNums)[0]
kf = pd.DataFrame(kf, columns = ['Credit Cycle Index'])

# ==============================================================================
# Fórmula de Vasicek para ajustar las TTC PD
# ==============================================================================
def vasicek_formula(ttc_pd, corr = 0.15, estado_ciclo = 0):
   adj_pd = norm.cdf((norm.ppf(ttc_pd) - estado_ciclo * math.sqrt(corr))/ math.sqrt(1 - corr))
   return adj_pd

# ==============================================================================
# Ajuste del vector de TTC PD de la matriz de transición
# ==============================================================================
# Únicamente la columna 'R' representa las probabilidades de default correspondientes a cada rating
de # la matriz. Extraemos esos valores para realizar el ajuste y obtener las PiT PD.
ttc_pd_vector = transition_matrix.iloc[:, -1] / 100
ttc_pd_df = pd.DataFrame(ttc_pd_vector)
ttc_pd_df = ttc_pd_df.rename(columns = {'R': 'PD'})

pd_adj = pd.DataFrame()

# Aplicación de la fórmula de Vasicek
for i in range(0, len(kf)):
   pd_adj['KF Period {0}'.format(i)] = vasicek_formula(ttc_pd_vector, estado_ciclo = kf.iloc[i][0])

# Gràfico con los ajustes posibles del vector
estados_sim = pd.DataFrame(np.arange(-3.0, 3.0, 0.01))
# Aquí generamos todos los estados posibles del índice del Ciclo del Crédito, que lo consideramos
como # una N(0, 1). Omitimos los valores que alejen dicha variable a más de 3 desvíos estándares por
# considerarlos no significativos.
pit_pd_vector = pd.DataFrame()
```

```
pit_pd_vector['Base'] = ttc_pd_vector

for i in range(0, len(estados_sim)):
   pit_pd_vector['Caso {0}'.format(i)] = vasicek_formula(ttc_pd_vector,
            estado_ciclo = estados_sim.iloc[i][0])

pit_pd_values = pit_pd_vector[pit_pd_vector['Base'] != 0]

# Preparamos los gráficos sólo para los valores de PD de la matriz que sean distintos de 0.
for i in range(0, len(pit_pd_values)):
   plt.plot(estados_sim, pit_pd_values.iloc[i, 1:], visible = 1,
         label = list(pit_pd_values.index.values)[i])#, 'r')

plt.title('PD como función convexa del Ciclo del Crédito')
plt.ylabel('Indice del Ciclo del Crédito (Z)')
plt.xlabel('PD como función de Z')
plt.legend(title = 'Rating')
plt.ion()
plt.show()
plt.ioff()

# ==========================================================================
# Simulación de PiT (TTC PD Ajustada por Convexidad)
# ==========================================================================
# Generamos variables aleatorias distribuidas uniformemente para simular probabilidades de default
sim_base = np.random.uniform(0, 1, 10000)
pd_sim = pd.DataFrame()
pd_sim['Base'] = sim_base

# Usamos la fórmula de Vasicek para ajustada cada una de las PD simuladas
# los estados posibles en base a una N(0, 1) con un Z entre -3 y 3
for i in range(0, len(estados_sim)):
   pd_sim['Caso {0}'.format(i)] = vasicek_formula(sim_base, estado_ciclo = estados_sim.iloc[i][0])

# Graficamos la distribución de probabilidad de cada PD en función del estado del ciclo, agrupándolas
en # rangos de 0.1 de probabilidad para poder ver los efectos que genera el ajuste para los diferentes
valores # de la PD
for i in range(0, len(estados_sim)):
   for j in range(1, 11):
      plt.subplot(5, 2, j)
      plt.subplots_adjust(hspace = 1.5, wspace = 0.3, left = 0.2, right = 0.9)
      h = j/10
      if ((j-1) * 0.1 < pd_sim.iloc[i, 0] <= j *0.1):
         plt.plot(estados_sim, pd_sim.iloc[i, 1:])
      plt.title('PD con Z entre ' + str(round(h - 0.1, 1)) + ' y ' + str(round(h,1)))
      plt.xlabel('Indice del Ciclo del Crédito (Z)')
      plt.ylabel('PD(Z)')

plt.ion()
plt.show()
plt.ioff()
```