



Universidad de San Andrés

Escuela de Administración y Negocios

Magister en Finanzas

Análisis de componentes principales: aplicación sobre la curva de rendimientos de bonos soberanos argentinos en moneda extranjera 2015-2019

Autor: Pablo Santiago Rodríguez Castro

DNI: 33.507.047

Director: Alejandro E. Loizaga

**Ciudad de Buenos Aires
Año 2021**

Resumen

Con cotizaciones de bonos soberanos argentinos en moneda extranjera correspondientes al período 2014-2019, se realizó un Análisis de Componentes Principales (PCA) para extraer de sus tasas de retorno diarias una descripción sintética de la dinámica de su estructura temporal. En línea con la literatura en la materia, se observó que el primer componente, asociado a movimientos paralelos en la curva de rendimientos, explicó el 89,0%, mientras que en conjunto con los factores vinculados a cambios en la pendiente y curvatura explicaron el 97,9% de la variabilidad de la curva durante el período seleccionado. La obtención y monitoreo de estos parámetros constituyen herramientas prácticas, tanto para las entidades del sistema financiero como para sus reguladores, tanto para reducir la dimensionalidad del análisis del movimiento en las curvas de rendimiento y facilitar su interpretación, como para medir la importancia relativa de estos componentes y de las distintas estrategias de cobertura frente al riesgo de mercado. A su vez, la disponibilidad actual de herramientas informáticas permite su cálculo y seguimiento de forma periódica y automatizada, como complemento para las métricas más comúnmente obtenidas y aquellas dispuestas por la normativa regulatoria.

Palabras clave: Componentes principales, reducción de dimensionalidad, curvas de rendimiento, valuación de activos financieros, estimación, predicción

Códigos JEL: C02, C51, C53, G12



Universidad de
San Andrés

1- Introducción	4
2- Marco teórico	6
3- Metodología	10
3.1- Los datos	11
3.2- Interpolación de curvas	13
3.3- PCA	19
4- Resultados	20
5- Conclusiones	25
6- Bibliografía	26
7- Anexo	27



Universidad de
SanAndrés

1- Introducción

En los últimos años, se ha observado que la gestión del riesgo de mercado ha adquirido un creciente protagonismo en el sistema financiero argentino. Tradicionalmente, entre las categorías de riesgo financiero, el análisis del riesgo de crédito siempre ha tenido un peso considerablemente superior por su dimensión determinante entre los factores explicativos de la rentabilidad de los bancos. Sin embargo, la mayor diversificación en las actividades de inversión de las entidades ha llevado a la dedicación de un mayor volumen de recursos y atención al estudio del riesgo vinculado a factores sistémicos, variables macroeconómicas y, particularmente, diversas tasas de interés, tanto en el análisis de la cartera de inversión como de la cartera de trading.

Existen numerosas métricas para las exposiciones que las entidades asumen frente al riesgo de mercado. Típicamente, la obtención del valor a riesgo de distintos activos o carteras ha sido el instrumento más difundido, tanto por su facilidad de cálculo e interpretación, así como por su gran contenido informativo. La *extreme value theory* ha complementado a este tipo de indicadores luego de que se los considerara insuficientes para la debida comprensión de los límites de la distribución de los retornos de activos, siendo que rara vez la misma se asemeja a una normal. Existe en la actualidad consenso en que la distribución de los retornos puede caracterizarse frecuentemente como leptocúrtica o de “colas pesadas”, lo que implica una mayor frecuencia de valores extremos (grandes pérdidas o grandes ganancias) respecto de una distribución normal. Particularmente en el estudio de las capacidades del sistema financiero para enfrentar situaciones severamente adversas y poco probables, esto ha generado la necesidad de continuar con la incorporación de nuevos instrumentos complementarios, entre los que se cuentan el cálculo de sensibilidades, diseño de escenarios, pruebas de estrés individuales, integrales, inversas, etc.

Dentro del herramental disponible, algunas métricas se enfocan en analizar las magnitudes probables de los cambios en el valor de un activo o una cartera en base a su volatilidad histórica, es decir, de la distribución histórica de sus retornos, sin necesariamente abordar el estudio de las relaciones entre los mismos y sus posibles *drivers*. Los indicadores que efectivamente se concentran en los factores explicativos que están por detrás de la volatilidad de los retornos, particularmente para los activos de renta fija, suelen vincularse al análisis del comportamiento de la curva de rendimiento que subyace al activo o cartera. Si bien la medición de sensibilidades frente a movimientos paralelos de la curva de rendimientos en la gestión de riesgo de mercado es de larga data, la incorporación de escenarios regulatorios que contengan movimientos no paralelos corresponde a los últimos años y resulta de utilidad tanto para la medición del riesgo de mercado de la cartera de trading como del riesgo de tasa de interés en la cartera de inversión (RTICI) de las entidades y las distintas subcarteras con las que los bancos puedan contar.

El análisis de componentes principales, si bien no es originario del análisis financiero, ha probado su utilidad en la reducción de la dimensionalidad de este problema: el de explicar, para una determinada clase de activo financiero, cómo varía en el tiempo la relación entre duración modificada (MD por su sigla en inglés¹), o madurez según el caso, y el riesgo percibido o rentabilidad exigida por los mercados (tasa interna de retorno o *yield to maturity*). Permite atribuir a unas pocas variables (construidas a su vez como combinaciones lineales de otras variables) los cambios en las estructuras temporales de tasas de interés para

¹ *Modified Duration*.

distintas clases y subclases de activos. Este tipo de análisis ha sido utilizado tanto para la construcción teórica de carteras de inversión óptimas de acuerdo al riesgo y retorno esperado de los activos que la componen y las interrelaciones de estos retornos, pero también en el campo específico de la gestión de riesgo.

Scheinkman y Litterman (Scheinkman y Litterman, 1991) popularizaron la aplicación de este método al análisis de curvas de rendimiento, concluyendo que la mayor parte de las variaciones de precios de los activos de renta fija puede descomponerse en tres factores, comúnmente conocidos como “nivel”, “pendiente” y “curvatura”. El método, en este caso particular, consistió en realizar transformaciones y combinaciones de las *yields* para distintas “*key maturities*”, para luego verificar cuáles de estas métricas resultaban más informativas de la volatilidad del conjunto. Un número reducido de los “componentes” obtenidos (combinaciones lineales de “*key maturities*”) podía usarse para dar explicaciones sintéticas e interpretables desde el sentido económico de la manera en que ocurren los cambios en las curvas de rendimientos a través del tiempo. Una vez sintetizados estos componentes, se puede proceder a investigar a qué factores (precios, tasas, valores de variables macroeconómicas) subyacentes a cada componente se debe prestar mayor atención.

El presente trabajo tiene como objetivo realizar un PCA sobre la curva soberana argentina en moneda extranjera para el período 2014-2019. La motivación del análisis, como se explicará a continuación, es la de determinar la proporción de la variabilidad de esta curva que puede ser explicada y pronosticada de manera precisa a partir de los 3 componentes clásicos ya en la literatura: los representativos del “nivel”, la “pendiente” y la “curvatura”. Los resultados servirán de base para futuras estimaciones de modelos predictivos, que podrán combinar tanto un enfoque de análisis de series temporales de cada componente, como también la bondad de ajuste del modelo a cada curva, en cada momento temporal.

Se procederá con la siguiente estructura: en una primera sección, se comentará el marco teórico, haciendo eje en una serie de estudios realizados sobre curvas soberanas de otros países, la información complementaria que se ha obtenido a partir de los distintos PCA y los resultados principales; en un apartado siguiente, se realizará una descripción precisa de la metodología PCA a aplicar, junto con los procedimientos para la obtención de las curvas que servirán de insumo para el modelo y una descripción de los datos utilizados; y para finalizar se hará un análisis de los resultados obtenidos y las conclusiones derivadas de los mismos.

2- Marco teórico

El Análisis de Componentes Principales (PCA, por sus siglas en inglés) fue introducido por Karl Pearson en 1901 y ha sido desde entonces reformulado y aplicado en un amplio espectro de disciplinas. Siendo una herramienta matemática capaz de reducir la complejidad en el análisis de cualquier fenómeno multidimensional, ha probado ser de utilidad en campos tan disímiles como la meteorología, la ingeniería y las finanzas. En cuanto a su aplicación específica al análisis de la estructura temporal de tasas de interés, Scheinkman y Litterman realizaron la primera aproximación al problema, la cual fue revisada e incrementada para diversos contextos en los últimos años en múltiples publicaciones. Los autores clasificaron a su análisis como un modelo de tres factores para explicar la variabilidad en las *yields* (YTM) de todos los activos financieros de renta fija en el mercado norteamericano. El hecho de que los factores “pendiente” y “curvatura” explicaran, junto al “nivel”, la casi totalidad de sus variaciones, permitiría a los inversores desarrollar estrategias de cobertura más eficientes que la de un portfolio de “cero duration” (que brindaba un *hedging* solo ante movimientos paralelos, o de nivel, en la estructura temporal).

Para su análisis, los autores entendieron que debían construir una “curva cero cupón implícita” para el mercado de *Treasuries* norteamericano. Esta sería una curva teórica, que no se correspondería exactamente con la que podría resultar de segmentar cada bono en cupones y capital, para luego valorar cada componente por separado (una curva cero cupón estricta), debido a la insuficiencia de instrumentos en el mercado. Más precisamente, los autores buscaron construir una curva, mediante precios de instrumentos cupón cero que, a cada fecha, mejor reflejen los precios observados de los bonos con cupón. Llamaron a estos retornos “*fitted yields*” o retornos ajustados y utilizaron para su análisis los retornos por encima de lo que consideraron una tasa libre de riesgo (la tasa genérica de *repo overnight*). Utilizando los tres factores mencionados, los autores encontraron que podían explicar el 96% de la variabilidad total de la curva y, mediante la descomposición de la varianza, adjudicaron un 89,5% promedio al factor “nivel”, un 8,5% al factor “pendiente” y un 2% al factor “curvatura”.

Tabla 1. Varianza explicada por factores PCA en cada punto fijo de madurez para la curva norteamericana (1986)

Madurez	Varianza total	Proporción de la varianza total explicada por		
		Factor 1	Factor 2	Factor 3
6 meses	99,5	79,5	17,2	3,3
1 año	99,4	89,7	10,1	0,2
2 años	98,2	93,4	2,4	4,2
5 años	98,8	98,2	1,1	0,7
8 años	98,7	95,4	4,6	0,0
10 años	98,8	92,9	6,9	0,2
14 años	98,4	86,2	11,5	2,2
18 años	95,3	80,5	14,3	5,2
Promedio	98,4	89,5	8,5	2,0

Fuente: Litterman y Scheinkman, 1991

Estos resultados llevaron a la deducción de que una cobertura que solo considera la MD (*duration-hedging*, la estrategia que se propusieron mejorar), efectivamente pudo prevenir las pérdidas derivadas del 90% de la volatilidad de la estructura temporal de tasas, para cualquier cartera de títulos soberanos, dentro del período analizado.

Numerosos autores retomaron la idea de Litterman y Scheinkman. Para el momento de su investigación, Diebold y Li (Diebold and Li, 2003) observaron que, si bien existían abundantes estudios respecto de la capacidad de pronosticar la variabilidad de la curva de rendimientos, la mayoría podían ubicarse en dos categorías notoriamente disímiles: aquellos que abordaban el problema desde el análisis puro de series temporales, aplicado principalmente sobre la tasa instantánea (“*equilibrium approach*”) y los que lo hacían desde el correcto ajuste estático de curvas momento a momento (“*arbitrage approach*”). En ambos casos, notaron que los resultados eran pobres en términos de capacidad predictiva. Los autores optaron por estimar los componentes exponenciales de Nelson-Siegel (1987) para poder así modelar la curva completa, periodo a periodo, con un único parámetro de tres dimensiones que evoluciona de manera dinámica, sintetizando ambas categorías de estudio existentes hasta el momento. A partir del cálculo de los componentes, estimaron modelos autorregresivos para cada uno, obteniendo resultados favorables en la predicción, particularmente para horizontes largos de inversión.

Un objetivo similar al de Diebold y Li fue perseguido por Delfiner (Delfiner, 2004). El autor realizó análisis segmentados sobre distintas ventanas temporales para diferentes tipos de activos argentinos, calculando variaciones semanales en los retornos pertenecientes al período inmediatamente anterior al default de 2001 (1999-2000) para Bontes, Letes y Bonos Globales. El primer obstáculo encontrado por el autor fue la falta de nuevas emisiones durante el período de análisis, lo cual no permitió trabajar con “*constant-maturity bonds*”. Según planteaba, el escenario ideal hubiera sido el de establecer puntos fijos de *maturity*, para luego obtener series temporales de tasas equivalentes, es decir, correspondientes a una misma *maturity* (algo virtualmente imposible de conseguir, por más líquido que sea el mercado). Dada la escasa emisión de nuevos títulos durante el período observado, la elección del autor fue la de tomar una ventana de análisis determinada (5 meses) y calcular la *maturity* promedio de cada bono observado a lo largo de esa ventana. Luego, se asignaron los cambios semanales en las YTM de cada bono al punto fijo de *maturity* correspondiente a dicho promedio. Esto permitió obtener varias series de retornos (diferencias entre tasas de rendimiento de un período y el subsiguiente, o ΔYTM), correspondientes a distintos puntos fijos de *maturity* (por ejemplo 6 meses y 1, 3, 5 y 10 años). Sobre este sistema multivariado Delfiner aplicó el análisis de componentes principales, obteniendo una estructura de varianzas y covarianzas entre las series de retornos para comprender su evolución conjunta o, lo que es lo mismo, la dinámica temporal de la curva de rendimiento subyacente. Si bien se plantea que una alternativa superadora a la de *maturities* promedio podría encontrarse en la inferencia de curvas de rendimiento a partir de alguna técnica, tal como el “*bootstrapping*”, nuevamente esto no resultaba posible por la escasez de bonos disponibles. Los resultados obtenidos por Delfiner indicaron que un 75% de la variabilidad total de la curva pudo atribuirse a cambios de nivel, un 15% a cambios de pendiente alcanzando, en promedio y junto con el factor curvatura, hasta un 96,6% para el subgrupo de Bontes y Letes.

Al igual que Delfiner, múltiples autores se han propuesto realizar este tipo de análisis. Los resultados en términos del poder explicativo que poseen los 3 componentes propuestos han presentado una relativa coincidencia a través de estos estudios, lo cual aporta una robustez mayor a esta metodología. Adicionalmente, los análisis complementarios que se han realizado a partir de los resultados son diversos, abarcando distintas temáticas dentro de los campos de la macroeconomía y las finanzas.

Nath y Golaka (Nath y Golaka, 2012) aplicaron el PCA al estudio de la estructura de tasas de interés de bonos soberanos de la India, para el período 1999-2012, aprovechando la digitalización y unificación de las operaciones financieras realizadas en el país a partir de fines de la década de 1990. Analizando la brecha entre tasas de corto y largo plazo, observaron cómo la importancia relativa de la pendiente de la curva de rendimientos se incrementó de manera significativa a lo largo de la crisis subprime. Durante dicho evento, si bien las tasas correspondientes a activos a 10 años de madurez permanecieron relativamente estables, las tasas de corto plazo se redujeron de forma considerable, más ligadas a la instrumentación de una política monetaria expansiva de baja de la tasa de Repo y encajes bancarios. Esto produjo que su análisis de componentes principales rindiera un resultado con una alta ponderación del factor “pendiente”, típicamente

el segundo componente principal. En el caso de Nath y Golaka, los movimientos en el nivel general, o movimientos paralelos de la curva de rendimientos, explicaron el 76,87% de su volatilidad, mientras que un 16,64% de la misma se explicó por cambios en la pendiente y un 5,14% de curvatura. La relativa importancia del segundo factor indicaría que una estrategia de cobertura en este mercado sería insuficiente si no incluyera instrumentos para compensar cambios repentinos en la pendiente de la curva, si bien en este caso lo que se observó fue particularmente un fuerte incremento en el valor de los activos de corto plazo debido a inyecciones de liquidez por parte de la autoridad monetaria. Los autores analizan, adicionalmente, el comportamiento en términos de series de tiempo de las variaciones diarias en las tasas de activos de distinta madurez, observando que las mismas tendían a revertir a su nivel medio y estaban autocorrelacionadas, aunque con una alta velocidad de decaimiento, mientras que su volatilidad registraba autocorrelación con mayores niveles de persistencia.

Cortés Espada et al (2008), por su parte, analizaron la dinámica de estructura temporal de tasas de activos soberanos mexicanos en moneda doméstica. Construyendo una curva cupón cero para el período 2002-2008, encontraron que el 95% de los movimientos de la curva podían ser explicados, en términos de PCA, por únicamente los dos primeros componentes (nivel y pendiente). Al mismo tiempo, encontraron una correlación positiva entre el factor “nivel” y las expectativas de inflación de largo plazo, así como una relación negativa entre el factor “pendiente” y la tasa overnight. Es decir que, al igual que en el estudio de Nath y Golaka, variaciones en la pendiente resultaron mayormente vinculadas a la volatilidad del tramo corto, registrando el tramo largo mayor estabilidad. Adicionalmente, comparan los resultados de un pronóstico realizado utilizando un modelo de factores (construido con los componentes principales obtenidos), respecto de las tasas observadas empíricamente, obteniendo un error medio de ajuste de tan solo 13 puntos básicos.

Ehrmann et al (2007) utilizaron el análisis de componentes principales para evaluar su hipótesis de convergencia entre las curvas de rendimiento de cuatro países europeos (Francia, España, Italia y Alemania) luego del establecimiento de la unión monetaria en la región. Para esto, como complemento del análisis de correlaciones entre curvas antes y después del evento, realizaron el mencionado PCA sobre una ventana de datos correspondientes al período 1993-1998 y, por otra parte, sobre la ventana 2002-2006. De esta manera, encontraron que luego del establecimiento de la unión monetaria, la reacción de los distintos países ante anuncios de política económica tendió a asemejarse, así como el peso de la volatilidad de las curvas que puede atribuirse a los 3 factores tradicionales en los PCA, mostrando de esta manera una convergencia significativa de los 4 países en términos de su exposición al riesgo de mercado.

Nuevamente respecto del caso mexicano, Mota Aragón y Mata Mata (2018) realizaron un PCA sobre la curva mexicana para un período considerablemente más amplio que el de la media de este tipo de estudios (1978-2017), encontrando una relación causal, en sentido de Granger, entre cada uno de los componentes principales calculados y variables macroeconómicas tales como la brecha del PBI, la tasa de inflación y la tasa de crecimiento económico.

Tabla 2 - Autovalores resultantes de un PCA sobre la curva mexicana (1978-2017)

Autovalores bajo PCA			
	1978-1988	1989-2008	2009-2017
	Variación acumulada	Variación acumulada	Variación acumulada
Componente 1	78,5%	52,7%	55,4%
Componente 2	98,4%	86,6%	86,4%
Componente 3	100,0%	96,6%	91,8%
Componente 4		98,7%	95,7%
Componente 5		99,7%	98,1%

Fuente: Mota Aragón y Mata Mata (2018)

En el análisis de causalidad de Granger, los autores encontraron que el PC1 suele causar incrementos posteriores en la inflación, lo cual, desde la perspectiva del estudio, sería razonable: ante mayores expectativas inflacionarias, el rendimiento que los agentes exigen de los activos de renta fija es mayor, lo cual genera un movimiento positivo de la curva soberana. Luego, siempre que las expectativas no sean erradas, se observa efectivamente el incremento en la tasa de inflación en mayor o menor medida. Por otra parte, el estudio ofrece evidencia para señalar que, en el caso mexicano, el PC2 (pendiente) ha explicado en el sentido de Granger posteriores caídas en el PBI. Nuevamente, los autores respaldan estos resultados con la extensa teoría económica que interpreta inversiones en la pendiente de una curva de rendimientos como alertas de una recesión inminente.

Por otra parte, Favero y Giglio (2006) estudian los efectos de la política fiscal sobre la curva de rendimientos italiana, en base a su spread con la curva alemana. El foco de su trabajo se encuentra en identificar de manera apropiada los efectos aislados de la política fiscal, sin que estos se vean contaminados por aquellos correspondientes al ciclo de negocios, cambios de régimen o impactos diferenciales de acuerdo a los tramos de *maturity* observados. Para esto, introducen una cadena de Markov, así como una variable dicotómica, para poder mantener el análisis antes y después de 1999, año en el que el Sistema Monetario Europeo se transformó en una Unión Monetaria. Al utilizar una cadena de Markov para modelar cambios de régimen, la información de expectativas de mercado de períodos sucesivos queda contenida en el estado “actual” de la variable, sin depender de estados anteriores. Luego, la distribución histórica de las variables observadas es utilizada para simular trayectorias futuras y analizar el efecto de los shocks. Esto permitió a los autores identificar que, durante la etapa previa a la adopción de una moneda única, diferenciales de inflación entre Italia y Alemania llevaban a las autoridades monetarias italianas a elevar la tasa de corto plazo, afectando el tramo corto de su curva soberana, algo que no se observa luego de adoptar ambos países una política monetaria única. Por otra parte, observaron que incrementos en el ratio deuda/PBI italiano sistemáticamente produjeron el efecto esperable de elevar el tramo largo de la curva soberana, dado el mayor riesgo percibido.

Con un objetivo análogo, Afonso y Martins (2010) han estudiado el efecto de shocks fiscales (cambios en el balance fiscal y en el ratio deuda/PBI) y monetarios (cambios en tasa de política) en Estados Unidos y Alemania. Su interés residía en identificar tanto shocks temporales como permanentes mediante funciones de impulso y respuesta. Para esto, obtuvieron los componentes principales de las respectivas curvas de rendimiento y analizaron su evolución con posterioridad a estos shocks. Su metodología consistió en una estimación por máxima verosimilitud de los parámetros de Nelson-Siegel, obtenidos en este caso utilizando filtros de Kalman. Luego, calibraron modelos VAR para proyectar los componentes (nivel, pendiente y curvatura) en relación con determinadas variables macroeconómicas (PBI real, inflación, tasa de política

monetaria), controlando la causalidad de los mismos en el sentido de Granger. Los autores encontraron que los shocks fiscales registraban un impacto relevante sobre la curva de rendimientos norteamericana, particularmente a través de la tasa de política, pero no sobre la curva alemana. Por otra parte, las curvas de ambos países sí se vieron afectadas por incrementos sorpresivos en los ratios de endeudamiento. El análisis de componentes de la curva de rendimientos permite, en este sentido, identificar variables macroeconómicas cuyo monitoreo, conjuntamente con la elaboración de escenarios, permiten realizar estrategias efectivas de cobertura. Los autores utilizaron información correspondiente al período 1980-2009, buscando incluir escenarios recesivos en ambas economías dado que estos constituyen la principal causa de necesidad de cobertura.

Nuevamente haciendo uso del método PCA para la realización de pronósticos, Suimon et al (2020) desarrollaron un modelo de aprendizaje automatizado basado en redes neuronales para, por un lado, realizar un análisis de componentes principales sobre la curva de rendimientos japonesa y, por otro, obtener un modelo predictivo que permita elaborar estrategias de *trading* en base a los resultados del análisis. La motivación de este trabajo consistía en que los modelos más popularmente utilizados para el análisis de curvas de rendimiento resultaban insuficientes para el estudio adecuado de la transición que se observó en la curva soberana japonesa hacia valores negativos entre 2015 y 2019, la cual sucedió bajo el contexto del Quantitative and Qualitative Easing implementado por el BoJ. Se mencionan los trabajos de Krauss et al (2017), Suimon (2018) y Matthew et al (2017) respecto de la utilización de redes neuronales para el análisis de mercados financieros, explicando que los modelos de alta complejidad, por ejemplo utilizando metodologías de *deep learning*, suelen toparse con dificultades vinculadas a su posterior interpretación. Se menciona además el estudio de Kikuchi y Shintani (2012), que proveen a los autores de una clasificación para los métodos de interpolación de curvas con la finalidad de valorar bonos soberanos en 4 categorías: polinomios por partes (un ejemplo popular es el de la utilización de *cubic splines*), no paramétricos, polinomios y funciones parsimoniosas (como la de Nelson-Siegel y Nelson-Siegel-Svensson). El método utilizado por Simon et al sería el de polinomios por partes, que brindarían los sets de *maturities* y tasas introducidos en el modelo *autoencoder*, el cual constituye otra forma de compresión de dimensionalidad de la volatilidad de una serie de variables, tal como lo es el PCA. En este caso, entre la capa de *inputs* y de *outputs* (en ambos casos, tasas correspondientes a distintas *maturities*), se ubica una capa oculta de 3 nodos, cuya interpretación resulta análoga a la de la mayoría de los modelos de PCA aplicados a curvas de rendimiento: nivel, pendiente y curvatura.

3- Metodología²

El presente trabajo construirá curvas de rendimiento para distintos puntos de MD fijos, establecidos sobre curvas resultantes de la interpolación de los pares de MD-YTM disponibles para cada ronda de operaciones, tanto por el método de Nelson-Siegel como por segmentación cúbica (CS, por su sigla en inglés)³. A partir de esta construcción, computando los valores de tasa para cada uno de estos puntos fijos, se realizará un PCA para determinar los principales drivers de su volatilidad. Se observará que el resultado brinda los componentes tradicionales interpretados como “nivel”, “pendiente” y “curvatura”, que a su vez coinciden con los factores exponenciales de Nelson-Siegel. Predecir estos 3 factores permitirá, entonces, pronosticar buena parte de la dinámica de esta curva. Qué proporción de esta dinámica quedará explicada por un modelo que pronostique de manera precisa la evolución conjunta de estos 3 factores, será la incógnita a resolver por el PCA, que a su vez garantizará por construcción la ortogonalidad de los factores.

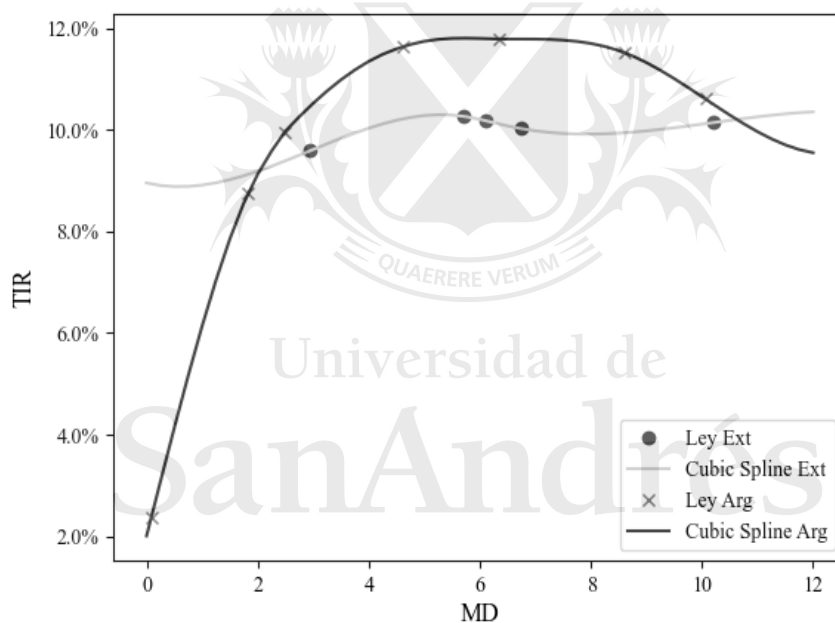
² El detalle del código utilizado para los cálculos contenidos en esta sección se encuentra desarrollado en el Anexo.

³ *Cubic spline*.

3.1- Los datos

Se obtuvieron las cotizaciones diarias correspondientes a BYMA⁴ para todos los bonos soberanos en moneda extranjera durante el quinquenio comprendido entre junio de 2014 y junio de 2019. Se evitó la inclusión de datos correspondientes al segundo semestre de 2019, debido a las complicaciones derivadas de los cambios en flujos de fondos vinculados a los sucesivos reperfilamientos de deuda realizados en ese momento por parte del gobierno nacional. Se procedió a eliminar bonos subsoberanos, para construir curvas con bonos uniformes en términos de riesgo de crédito y se tomaron, siempre que hubiera más de un precio disponible, los promedios de las cotizaciones con plazo de liquidación inmediata, a 24, 48 y 72hs. Luego, observando la separación durante amplios segmentos de tiempo de las cotizaciones de bonos Ley Argentina y Ley Nueva York, se optó por mantener únicamente los primeros, dado su mayor volumen operado⁵ y considerando que para la construcción de una curva de rendimientos es necesaria una familia homogénea en términos del riesgo y rendimiento exigidos por el mercado.

Gráfico 1. Comparación de curvas de rendimiento Ley Argentina y Ley Extranjera



Fuente: elaboración propia en base a cotizaciones correspondientes al día 30 de octubre de 2018, obtenidas de BYMA para los siguientes activos: AN18D, AO20D, AY24D, PARAD, DICAD, AA25D y AA37D (Ley Argentina) y A2E2D, A2E7D, A2E8D, AA26D y AA46D (Ley Extranjera).

Posteriormente, se realizó un recorte en los extremos de la ventana temporal para acabar con un período para el cual se cuente, tanto al principio como al final, con suficientes bonos cotizantes por día para poder interpolar una curva de rendimientos, siendo la ventana definitiva para la realización del PCA aquella comprendida entre el 13 de noviembre de 2015 y el 1 de febrero de 2019. Este recorte responde a la necesidad del método de interpolación por CS de al menos 5 puntos de construcción para ajustar una curva. Se optó, para obtener un mejor resultado en el ajuste de las curvas, incorporar a la ventana de análisis

⁴ Bolsas y Mercados Argentinos

⁵ La diferenciación entre “Ley Argentina” o “Ley Nueva York” forma parte de las condiciones de emisión de estos títulos, indicando que ante cualquier diferendo entre las partes, las mismas se resolverán en uno u otro tribunal. La separación de las curvas de rendimientos correspondientes a ambos subtipos de activos resultó notoria durante el segundo semestre de 2018.

únicamente ruedas con cotizaciones para al menos 6 bonos, dentro de los 8 activos Ley Argentina disponibles. El resultado fue un set de cotizaciones promedio diarias correspondientes a la ventana indicada para los siguientes 8 bonos en dólares, Ley Argentina: AA17D, AN18D, AO20D, AY24D, PARAD, DICAD, AA25D y AA37D.

Una vez obtenidas las series de precios, se procedió a detallar los flujos de fondos para cada activo, de manera de poder calcular para cada fecha y cada activo su tasa interna de retorno. El cálculo se realizó restando al precio actual los flujos futuros descontados a una tasa incógnita, igualando a cero esta diferencia e iterando mediante el método de Newton-Raphson⁶, consistente en la siguiente expresión:

$$X_{n+1} = X_n - \frac{f(X_n)}{f'(X_n)} \quad f'(X_n) \neq 0$$

El cálculo del valor presente de cada activo se hizo para una convención de 365 días al año, la cual se corresponde con todos los activos seleccionados.

Tabla 3. Estadísticas descriptivas de los precios utilizados y tasas obtenidas⁷

	TIR	TTM	MD
Nro. Obs.	5.951	5.951	5.951
Media	7,2%	3.568,0	5,3
Desvío	3,0%	2.748,8	3,5
Mínimo	-3,5%	3,0	0,0
25%	4,9%	1.007,0	2,4
50%	7,1%	2.698,0	4,5
75%	9,2%	6.500,5	7,9
Máximo	17,0%	8.848,0	13,0

Fuente: elaboración propia en base a cotizaciones correspondientes al período del 1 de enero de 2014 al 1 de febrero de 2019, obtenidas de BYMA para los siguientes activos: AA17D, AN18D, AO20D, AY24D, PARAD, DICAD, AA25D y AA37D.

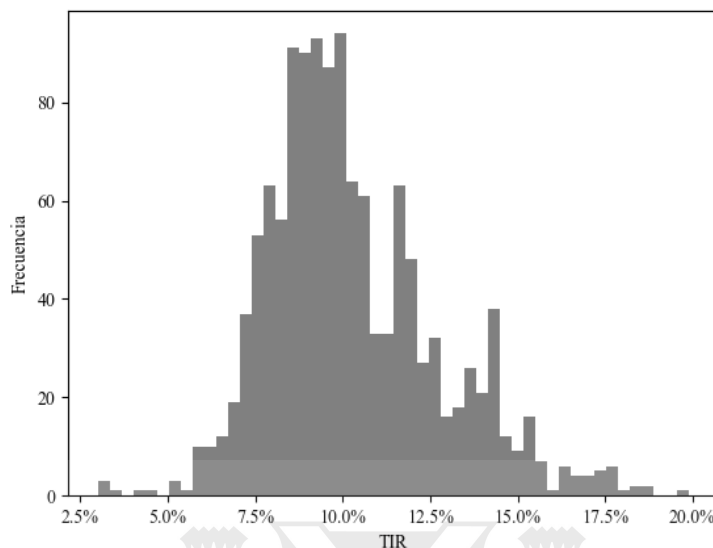
Se observa en las estadísticas descriptivas de los bonos utilizados para la construcción de las curvas diarias un *Time to Maturity* mínimo de 3 días (MD de 0,00874 años), correspondiente tanto al AA17D como al AN18D que alcanzaron sus vencimientos dentro de la ventana de análisis⁸. En el otro extremo, el valor máximo de 13 años de MD corresponde al bono PARAD al inicio de la ventana temporal. Respecto de la distribución de los rendimientos, se observa un valor medio razonable para el período, de 7,2%, con un desvío estándar de 3,0 pp. Por otra parte, los valores extremos, tanto el valor negativo mínimo como el valor máximo, se corresponden con rendimientos calculados para el AA17D y AN18D cercanos a sus respectivos vencimientos, cuando las tasas de los activos pueden registrar este tipo de irregularidades. Para mejorar el ajuste de curvas, se procedió a eliminar este tipo de *outliers*.

⁶ Puede observarse el código utilizado para el cálculo en el anexo.

⁷ Tasa interna de retorno (TIR), Madurez (TTM), Duration Modificada (MD).

⁸ Cotizaciones correspondientes a 1 y 2 días de TTM fueron excluidas por registrar valores de TIR muy lejanos a la distribución general, consideradas *outliers*.

Gráfico 2. Histograma de los rendimientos obtenidos



Fuente: elaboración propia en base a 5.951 cotizaciones correspondientes al período del 1 de enero de 2014 al 1 de febrero de 2019, obtenidas de BYMA para los siguientes activos: AA17D, AN18D, AO20D, AY24D, PARAD, DICAD, AA25D y AA37D.

Todas las curvas fueron calculadas para un marco temporal fijo, dentro del cual se establecieron puntos de análisis. Para todos los días de la ventana se calculó la TIR correspondiente a una MD de, por ejemplo, un año. De esta manera, analizar la evolución de la TIR correspondiente a este punto implica examinar la evolución de la altura de la curva para un valor de MD de un año. Hacer lo mismo para 48 puntos fijos (demarcados por 12 años divididos en trimestres) nos permite estudiar la movilidad de toda la curva, entre el punto mínimo de MD establecido y el máximo, día tras día, para toda la ventana temporal entre noviembre de 2015 y febrero de 2019. Se optó por fijar el máximo en 12 años, coincidiendo con la MD promedio del bono de vencimiento más lejano (PARAD).

3.2- Interpolación de curvas

Dado que las cotizaciones de los bonos disponibles no siempre coinciden en MD con los puntos fijos establecidos y necesarios para una comparación discreta período a período, se recurre a la interpolación por *cubic splines* (CS). Esto es una función no lineal, construida en base a polinomios de tercer grado. Los mismos deben pasar por los puntos de control establecidos por los pares de MD-TIR obtenidos de las cotizaciones vigentes y contar con primeras y segundas derivadas continuas para todo el segmento temporal analizado. Se establece una derivada segunda igual a cero en los extremos de las curvas como restricción para poder dar solución al sistema de ecuaciones, lo cual produce pobres resultados al utilizar esta metodología para extrapolar (la pendiente de cada polinomio se vuelve constante en los extremos). Es por esta razón que se limitó la ventana temporal en puntos cercanos a los extremos de la distribución de los datos disponibles (como se explicó en el apartado anterior, 0 y 12 años).

Para $n+1$ puntos (y_0, y_1, \dots, y_n) , el fragmento número “ i ” del *spline* (segmento) puede representarse de la siguiente manera⁹:

$$Y_i(t) = a_i + b_i t + c_i t^2 + d_i t^3,$$

donde “ t ” es un parámetro entre 0 y 1, e $i=0, \dots, n-1$, entonces

$$Y_i(0) = y_i = a_i,$$

$$Y_i(1) = y_{i+1} = a_i + b_i + c_i + d_i,$$

derivando $y_i(t)$ en cada intervalo resulta

$$Y_i'(0) = D_i = b_i,$$

$$Y_i'(1) = D_{i+1} = b_i + 2c_i + 3d_i.$$

Resolviendo 2 a 4

$$y_i = a_i,$$

$$b_i \equiv D_i,$$

$$c_i = 3(y_{i+1} - y_i) - 2D_i - D_{i+1},$$

$$d_i = 2(y_i - y_{i+1}) + D_i + D_{i+1},$$

y exigimos que la segunda derivada también coincida en los puntos de control

$$Y_{i-1}(1) = y_i,$$

$$Y_{i-1}'(1) = Y_i'(0),$$

$$Y_i(0) = y_i,$$

$$Y_{i-1}''(1) = Y_i''(0).$$

Dado que los puntos interiores y extremos satisfacen

$$Y_0(0) = y_0,$$

$$Y_{n-1}(1) = y_n,$$

dando un total de $4(n-1)+2=4n-2$ ecuaciones para $4n$ incógnitas. Se establecen para posibilitar la resolución dos condiciones adicionales, se requiere que la segunda derivada de los extremos sea igual a cero:

$$Y_0''(0) = 0,$$

$$Y_{n-1}''(1) = 0,$$

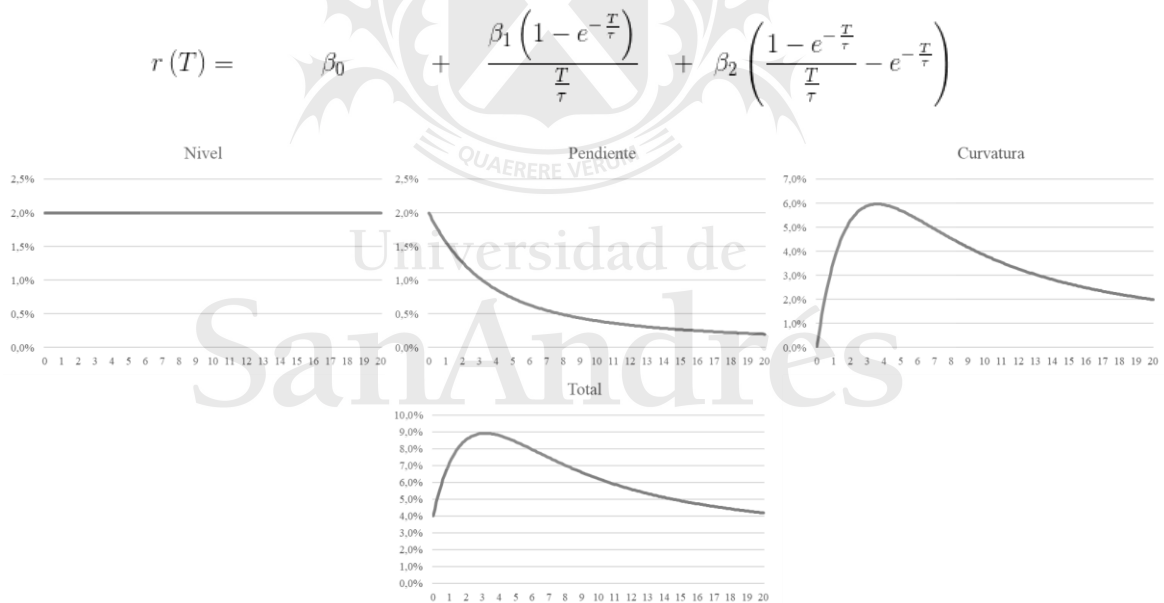
brindando el siguiente sistema de ecuaciones

⁹ Para mayor detalle puede consultarse <https://mathworld.wolfram.com/CubicSpline.html>

$$\begin{bmatrix} 2 & 1 & & & & & \\ 1 & 4 & 1 & & & & \\ & 1 & 4 & 1 & & & \\ & & & 1 & 4 & 1 & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & 1 & 4 & 1 \\ & & & & & 1 & 2 \end{bmatrix} \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_{n-1} \\ D_n \end{bmatrix} = \begin{bmatrix} 3(y_1 - y_0) \\ 3(y_2 - y_0) \\ 3(y_3 - y_1) \\ \vdots \\ 3(y_{n-1} - y_{n-3}) \\ 3(y_n - y_{n-2}) \\ 3(y_n - y_{n-1}) \end{bmatrix}.$$

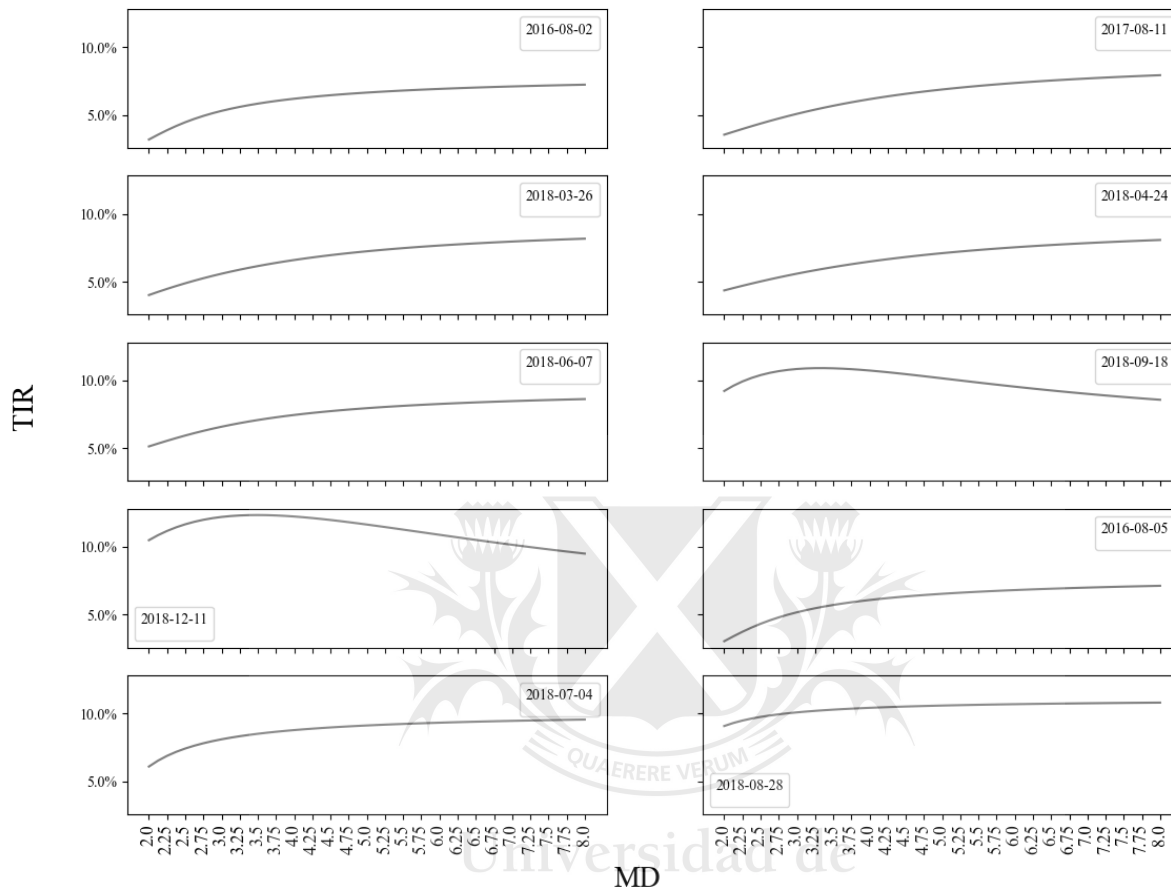
Esta metodología permite contar con una estimación continua de la curva de rendimientos más suavizada que una interpolación lineal. Para el análisis de componentes principales, obtener series de retornos de curvas interpoladas vía el método más comúnmente utilizado de Nelson-Siegel presentaría un sesgo que favorecería el resultado que se espera obtener. Esto es debido a que la estructura de la interpolación permite variar tres factores de construcción únicamente en tres orientaciones que coinciden con las características de nivel, pendiente y curvatura. Esto produce que el resultado de un PCA sobre curvas construidas mediante esta metodología siempre sea del 100%, es decir, los tres primeros componentes principales explicarían el 100% de la variabilidad de las curvas. Esto queda en evidencia al exponer la fórmula para la interpolación propuesta por Nelson y Siegel:

Gráfico 3. Contribución de los factores en la interpolación por Nelson Siegel



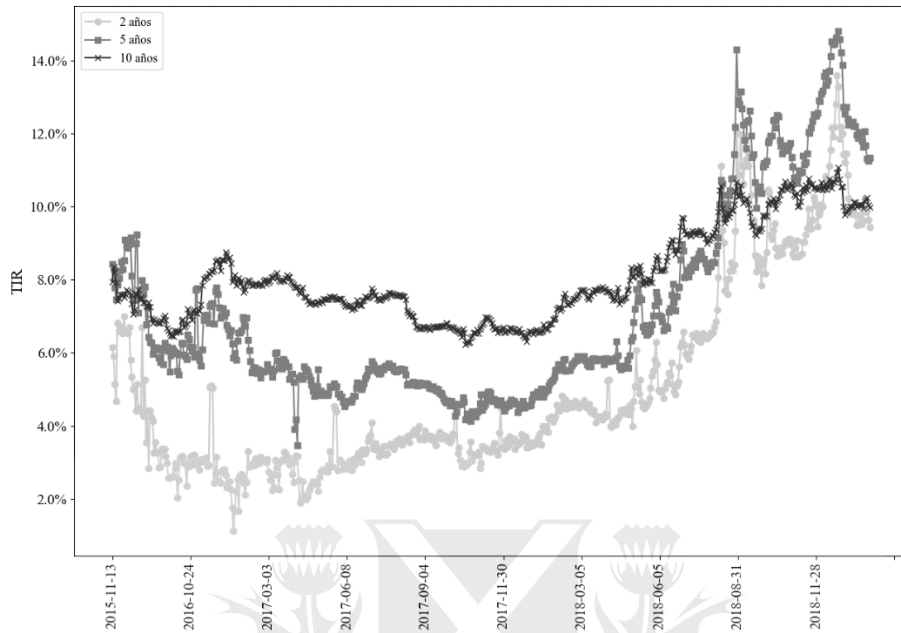
A modo de ejercicio, se realizó el PCA, obteniendo las series de TIR correspondientes a los puntos fijos de MD interpolando las curvas diarias mediante esta metodología, presentando los siguientes resultados:

Gráfico 4. Muestra aleatoria de curvas diarias interpoladas por Nelson Siegel



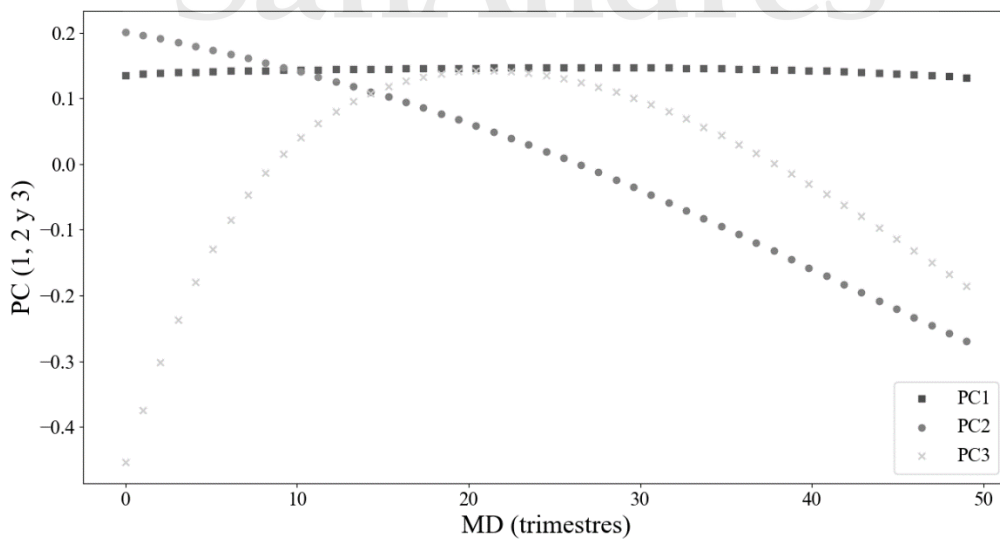
A partir de las curvas interpoladas, se obtuvieron series de tasas correspondientes a los 48 puntos de MD de análisis. En el siguiente gráfico se expone la evolución de los valores de tasa correspondientes a los 2, 5 y 10 años de MD, para toda la ventana de análisis:

Gráfico 5. Evolución de las YTM correspondientes a 2, 5 y 10 años de MD



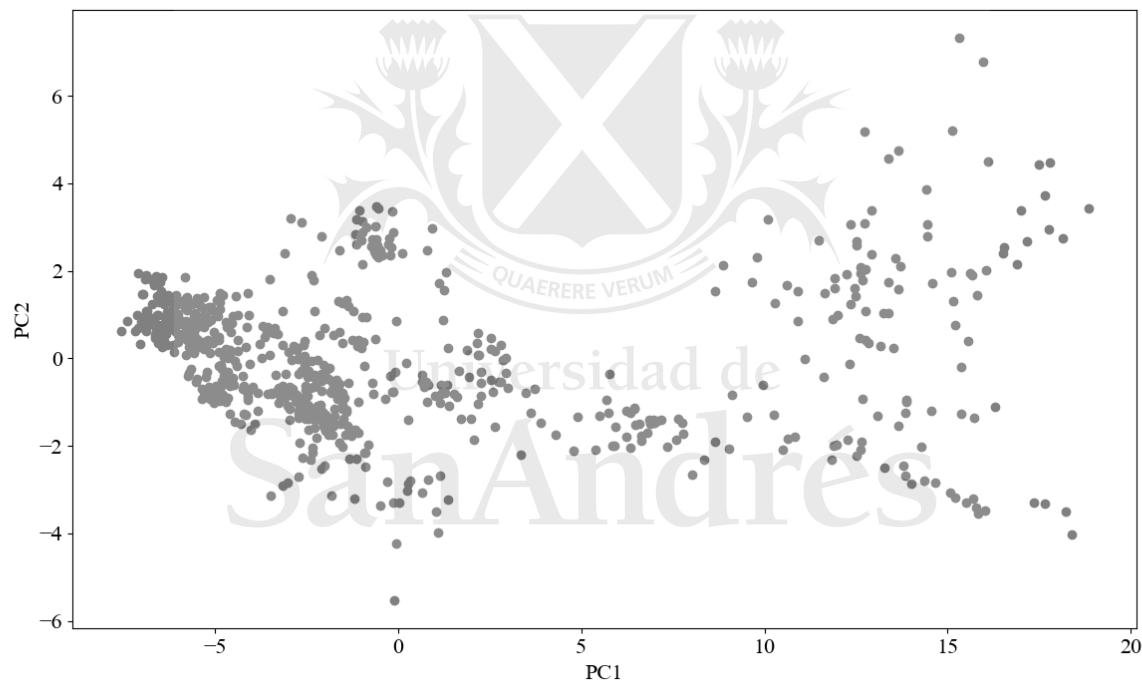
Finalmente, se utilizaron las series de tasas obtenidas para calcular la matriz de varianzas y covarianzas entre las mismas, luego de su estandarización, y se aplicó el modelo de PCA cuyos detalles serán desarrollados en el próximo apartado. Se optó por sintetizar 3 componentes principales a partir de los datos, por lo que se obtuvo una matriz de 3 autovectores, contruidos a partir de las 48 series de tasas, que constituyen según la metodología utilizada la forma más precisa de representar el dataset original en 3 dimensiones.

Gráfico 6. Construcción de los tres componentes principales



Cada componente principal puede entenderse como una nueva variable, construida como una combinación lineal de los valores de las variables originales (en nuestro caso, valores de tasa estandarizados, correspondientes a cada uno de los 48 puntos de MD establecidos). En el gráfico 6 podemos observar la correlación de cada componente con las tasas registradas en cada uno de estos puntos. Respecto del PC1, vemos que es una combinación que correlaciona positivamente con valores en todo el espectro de MD (o lo que es equivalente, asigna coeficientes positivos a todos los valores de MD). Es decir que las variaciones de la curva que se encuentran contenidas en este componente, son movimientos de tasa en el mismo sentido y de similar magnitud para todo punto de MD o, dicho de otra manera, movimientos paralelos de la curva. Por otro lado, el PC2 correlaciona en sentido positivo para los niveles de tasas correspondientes a entre 0 y 25 trimestres de MD, y en sentido opuesto para niveles de 26 trimestres de MD en adelante. Es decir que este componente condensa los cambios en la pendiente de la curva. Finalmente, el tercer componente presenta movimientos de la curva en igual sentido en ambos extremos, y de sentido contrario para los puntos intermedios de MD, es decir, cambios en la curvatura. Por construcción, los tres componentes son ortogonales, por lo que la volatilidad explicada por cada uno no está contenida en ninguno de los otros dos.

Gráfico 7. Distribución conjunta de los componentes 1 y 2



Los valores diarios de tasa correspondientes a los distintos puntos de MD de la curva de rendimientos pueden expresarse a través de los primeros dos componentes principales. En este caso, vemos que la disposición de los datos se ve concentrada mayormente a lo largo del eje horizontal, representativo del PC1, respecto del eje vertical (PC2). Debe destacarse que, para poner el foco en cambios relativos en los valores de tasa, cada serie fue estandarizada, restando su media y dividiendo por su desvío estándar. De esta manera, vemos que son más los casos en los que la distribución se encuentra por encima de los 15 desvíos estándar para la variable PC1, que para PC2, por lo que la primera explica una mayor proporción de la dispersión del set de datos original. Puede agregarse, al observar el gráfico, que los mayores cambios positivos en el nivel de la curva (que en la muestra se observan a partir de mediados de 2018), a la derecha del gráfico 7, coinciden con una mayor dispersión en los movimientos de la pendiente, la cual de hecho se invierte durante el mismo período.

El porcentaje de la variabilidad total de la curva de rendimientos que es explicado por estos 3 factores es del 99,95%, siendo 94,12% para el PC1, 5,32% para el PC2, y 0,05% para el PC3. Estos valores anormalmente altos se deben justamente a la forma de construcción de las curvas por interpolación de Nelson-Siegel en base a 3 factores que, como se expuso, permiten covariar a los puntos de la curva en estos mismos 3 sentidos. En la próxima sección se profundizará en la metodología de cálculo de los componentes y en una sección posterior, se brindarán los resultados del PCA para curvas interpoladas por un método alternativo, que no asuma el comportamiento ya explicado en el método de Nelson-Siegel.

3.3- PCA

El análisis de componentes principales es un método ampliamente utilizado al estudiar problemas que requieren interpretaciones simples de múltiples características correlacionadas de un mismo fenómeno. Este es el caso de la curva de rendimientos, para la cual los rendimientos (YTM) correspondientes a cada plazo de inversión (medida como TTM o MD) varían en conjunto de acuerdo a patrones complejos. Una matriz de varianzas y covarianzas, o de correlaciones, perteneciente a series de variaciones de las YTM para cada MD (dependientes de las sucesivas cotizaciones disponibles), conforma un fenómeno difícil de interpretar de manera exhaustiva. Sin embargo, al seleccionar un número finito de puntos representativos del horizonte temporal comprendido en la curva soberana, podemos limitar una de las dimensiones de análisis.

En el presente caso, la muestra de bonos en moneda extranjera con suficiente liquidez, permitió construir para cada fecha en la ventana de estudio una curva interpolada, para la cual los puntos de construcción disponibles en base a las combinaciones de YTM y MD resultantes de sus cotizaciones y flujos de pagos se ubicaron en todos los casos entre cero y doce años. Dado que estas combinaciones son variables y por ende no comparables unas con otras, la interpolación mediante el CS probó ser un método eficaz para obtener, para un set de 48 MD preestablecidos como “clave” (en un sentido similar al del análisis por *key rate durations*), series de YTM. De esta manera, luego de establecer los flujos de pagos para todos los activos de la muestra, calculado las YTM diarias para cada activo en cada fecha en el período 2015-2019, e interpolado una curva continua a partir de estos pares ordenados, se procedió a construir una serie de YTM pertenecientes a los valores fijos representativos de MD: cada límite trimestral entre el año 0 y el 12. Al calcular los valores diarios resultantes de estas 48 series de tasas de rendimiento, se evaluó su variabilidad conjunta mediante el método PCA.

El mismo consiste en la obtención de autovalores y autovectores de su matriz de varianzas. La matriz de autovectores obtenida puede entenderse como una serie de vectores ortonormales que constituyen combinaciones lineales de las series originales utilizadas como insumo. De esta manera, el patrón de variabilidad explicado en un primer componente principal, no coincidirá con la variabilidad contenida en los sucesivos componentes de orden mayor.

Para comenzar, partimos de una matriz “A” simétrica de dimensiones $n \times n$ y rango $r=n$, en nuestro caso, la matriz de varianzas y covarianzas de las “n” (48) series de YTM. El hecho de que la matriz sea simétrica, significa que la misma se puede ortogonalizar:

$A = V\Lambda V^T$ donde Λ es una matriz diagonal que contiene los “r” autovalores λ_i de A:

$$\lambda_i = \text{diag}[\lambda_i]$$

Y V es una matriz ortogonal, por lo que su traspuesta es igual a su inversa

$V^T = V^{-1}$ que a su vez está compuesta por “r” autovectores “ v_i ” ortonormales ($v_i^T v_j = 0$ y $v_i^T v_i = 1$, para todo $i \neq j$).

Además

$$A v_i = \lambda_i v_i \quad i=1..,r$$

Es decir que, al operar la matriz A sobre el autovector “i”, el resultado es el mismo autovector, multiplicado por una constante, que es su correspondiente autovalor “i”. En nuestro caso particular, podremos obtener nuestras variables transformadas, o componentes principales “ x_i ”, a partir de nuestros valores diarios de tasa “ y_i ” de la siguiente manera

$$x_i = \sum_{j=1,n} (v_{ij})^T y_j$$

Puesto de otra forma, al buscar el componente principal de una matriz de varianzas y covarianzas “A”, buscamos aquel vector de módulo unitario “ v ” que maximice $v^T A v$. Este será el vector que mejor refleje la variabilidad contenida en “A” y, al mismo tiempo, que minimice la suma de las proyecciones de cada punto de la muestra sobre un vector de la dimensión de v .

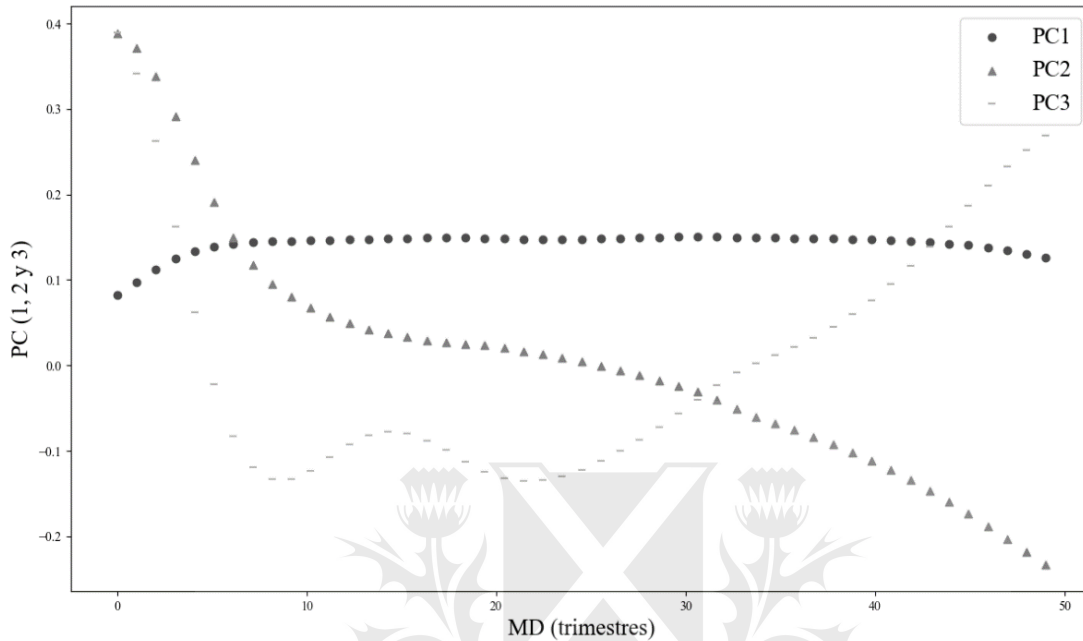
El método PCA permite convertir una matriz de varianzas y covarianzas en, por un lado, “escala” (dada por los autovalores) y, por otro, “dirección” (dada por los autovectores). Al sumar la totalidad de direcciones en las que la información original varía, escaladas por sus correspondientes autovalores, podemos recuperar la totalidad de la matriz original. Alternativamente, podemos mantener solamente algunos segmentos de interés y cumplir con la función primaria de este tipo de análisis: reducción de dimensionalidad. El cociente entre cada autovalor y la sumatoria del total de autovalores, será equivalente al aporte de varianza del componente principal asociado a dicho autovalor a la varianza total del set de datos.

4- Resultados

Analizando los resultados para 5 años de curvas interpoladas por CS, observamos que la proporción de la varianza total de las 48 series de pares de MD fija y TIR que resulta explicada por los 3 componentes principales es del 97,89%. Para analizar únicamente variaciones relativas, dado que los niveles de los componentes pueden tener escalas muy diferentes y eso dificulta su análisis conjunto, se procedió a estandarizar cada variable, sustrayendo su media y dividiendo por su desvío estándar.

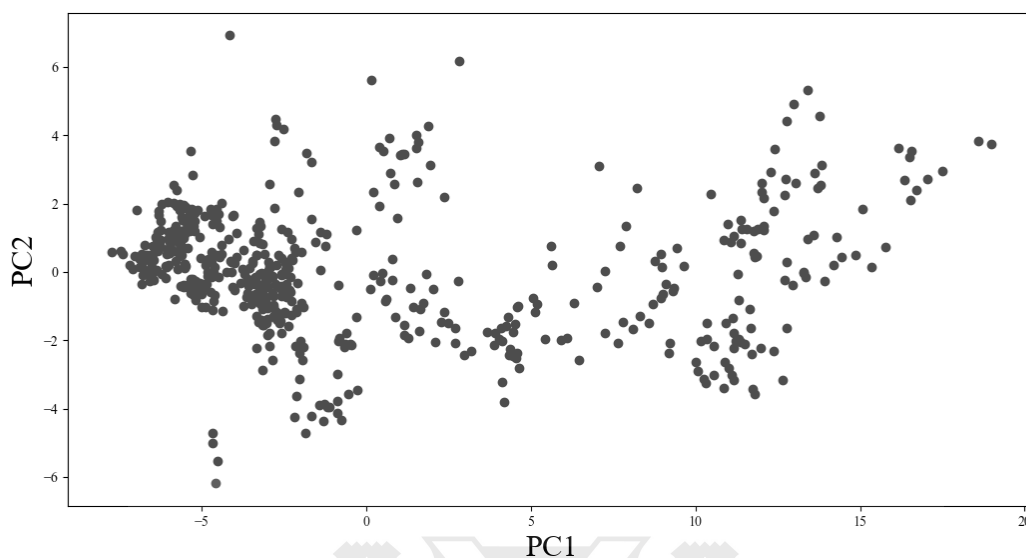
La disminución del poder explicativo producida por el cambio en el método de interpolación (de N-S a *cubic spline*) fue de solamente 2,09%. Descomponiendo entre los 3 primeros componentes ortonormales, se observa que el primero explica el 89,00% de la varianza, el segundo un 6,68% y el tercero un 2,21%.

Gráfico 8. Construcción de los tres componentes principales



Puede observarse que las correlaciones de cada componente con las YTM a lo largo del espectro de MD son análogas a aquellas obtenidas con curvas interpoladas por N-S: el primer componente es una combinación lineal relacionada positivamente con todos los valores de MD, que podemos asociar al nivel. El segundo pondera positivamente los primeros valores y negativamente aquellos superiores a los 25 trimestres de MD. Adicionalmente, resulta aparente que la mayor variabilidad asociada a cambios de pendiente se produce en las tasas de corto plazo (0 a 10 trimestres). Finalmente, el tercer factor principal pondera positivamente MD inferiores a los 5 trimestres y superiores a los 35 trimestres (extremos) y negativamente a los valores intermedios, representando nuevamente la curvatura.

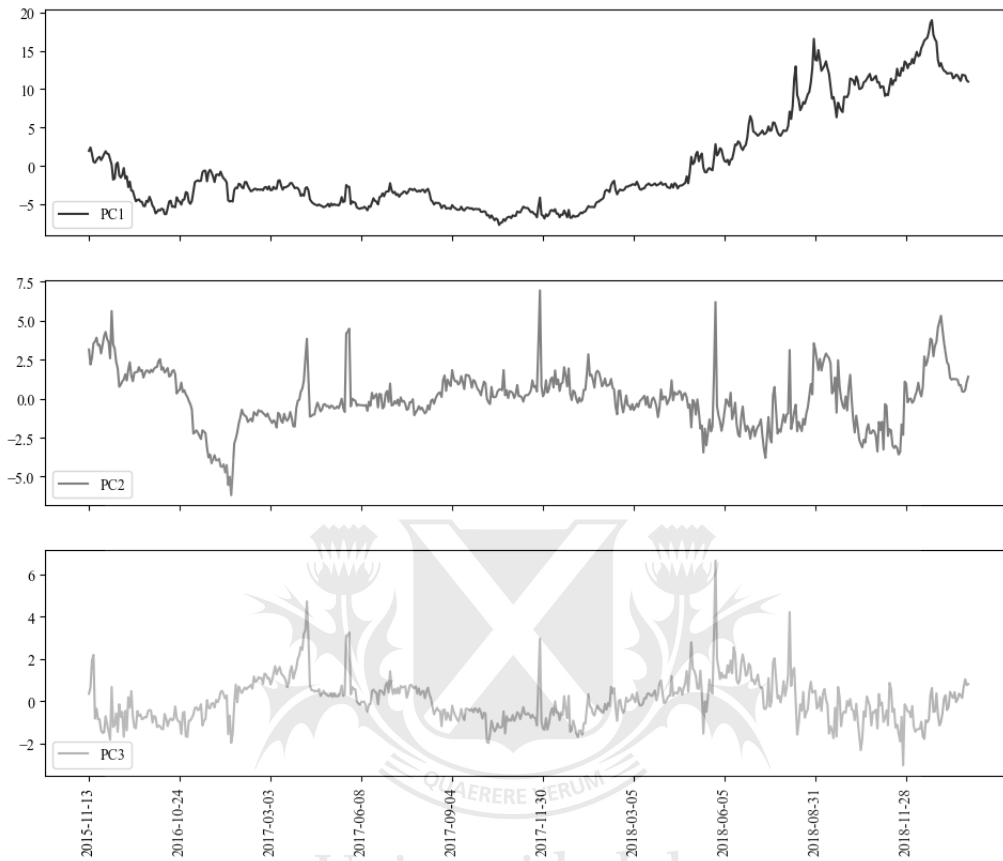
Gráfico 9. Representación de los datos en base a los dos primeros componentes principales



Sabiendo que los 2 primeros PC explican el 95,68%, es posible representar prácticamente la totalidad de la información contenida en las 48 series de retornos diarios para una ventana de 5 años, en función de las dos nuevas variables estandarizadas obtenidas, que simbolizan cambios en el nivel y en la pendiente de la curva de rendimientos. Vemos que los cambios proporcionales al desvío estándar de cada una de las variables compuestas son considerablemente mayores para el PC1 que para el PC2, cuyos valores para toda la muestra no superan en valor absoluto los 6 desvíos estándar.

Respecto de la evolución de los ponderadores asociados a cada componente, podemos ver que la evolución del nivel de la curva presentó una tendencia negativa hasta mayo de 2018, momento en el que comienza a existir incertidumbre respecto de la sostenibilidad de la deuda en moneda extranjera, coincidiendo con una aceleración en el tipo de cambio que comenzó con una depreciación mensual de 25% y que alcanzaría el 100% un semestre después. Valores negativos en el ponderador de la pendiente implican un impacto relativamente negativo sobre las tasas de corto plazo y positivo sobre las de largo, por lo que se observa hacia fines de 2016 y principios de 2017 es una reducción en las tasas de corto en relación a las de largo. La mencionada volatilidad en la curva a partir de mediados de 2018 se tradujo no solo en incrementos de nivel, sino también como puede verse en la evolución del PC2, incrementos y reducciones de la pendiente. Respecto de la curvatura, vemos que los valores extremos de la curva aumentaron en relación a los valores medios tanto al principio como al final de la ventana temporal, sucediendo un moderado aplanamiento entre mediados de 2017 y de 2018.

Gráfico 10. Evolución de los ponderadores de cada componente principal

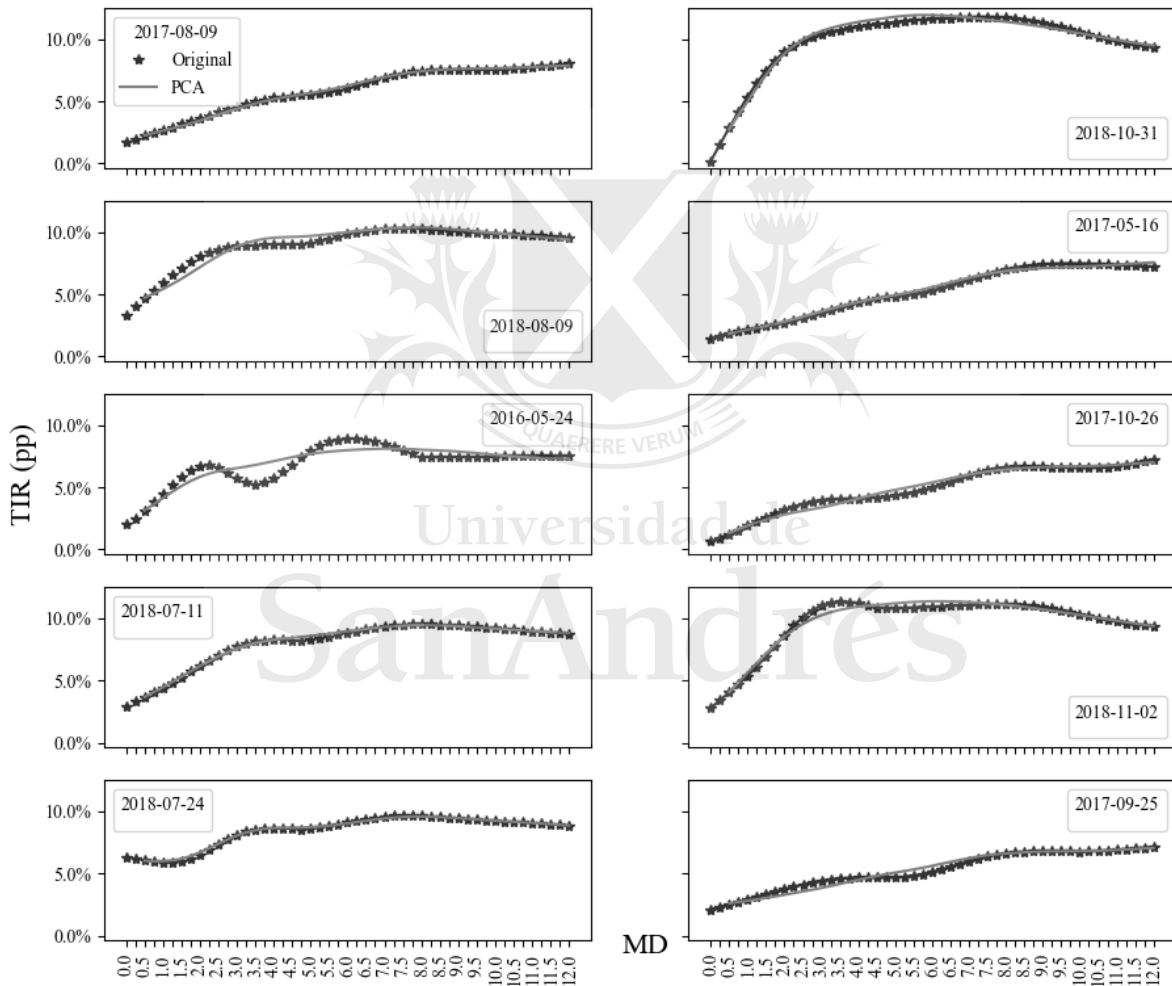


Por último, además de validar las conclusiones a menudo encontradas en la bibliografía respecto de los tres factores clásicos que explican la casi totalidad del comportamiento temporal de las curvas, podemos verificar una utilidad adicional de la reducción de utilidad como herramienta. Es posible reconstruir, utilizando los autovectores de la matriz de varianzas y covarianzas generada con los valores originales de TIR-MD, la información original. Es decir que, partiendo de una matriz de 582 observaciones de valores de TIR para 49 puntos fijos de MD, obtuvimos una matriz de varianzas y covarianzas de 582×582 , de la cual pudimos sintetizar 2 nuevas matrices de menor tamaño, a partir de las cuales podemos replicar la información original con mínima pérdida. Estas son, por un lado, la matriz de autovectores ortonormales (3×48) que sintetizan la matriz de varianzas y covarianzas en los 3 componentes de mayor orden que hemos mantenido y, por otro, la matriz de pesos o ponderadores adjudicados en cada día de la ventana temporal a cada uno de los 3 componentes (582×3). De manera similar, los resultados encontrados indican que, de construirse 3 variables simples, representativas de los tres componentes obtenidos, las mismas podrían brindar una buena explicación de la dinámica temporal de la curva en cuestión¹⁰. Dicha explicación contará, adicionalmente, con una interpretación económica que estará anclada en únicamente los componentes seleccionados, nuevamente, con mínima pérdida de información.

¹⁰ Dicha construcción no debe necesariamente ser el resultado de un análisis PCA. Distintos autores computan la variable pendiente como la diferencia entre las yields correspondiente a 10 años y a 1 año. Como medida de la curvatura, puede utilizarse la diferencia entre la sumatoria de las yields a 1 y 10 años y dos veces la yield para una MD de 5 años. Un promedio entre estas tasas puede servir de proxy para el nivel.

Tal como se ha observado en numerosas investigaciones, análisis de series temporales sobre los componentes obtenidos pueden ser de utilidad para predecir la evolución de la curva de rendimientos, así como para la realización de simulaciones para la construcción de escenarios de estrés y otras métricas de apoyo en la gestión de las exposiciones al riesgo de mercado. A continuación pueden observarse las curvas diarias reconstruidas a partir de los 3 componentes obtenidos en el presente análisis, en comparación con las curvas originales, para observar la precisión de la estimación:

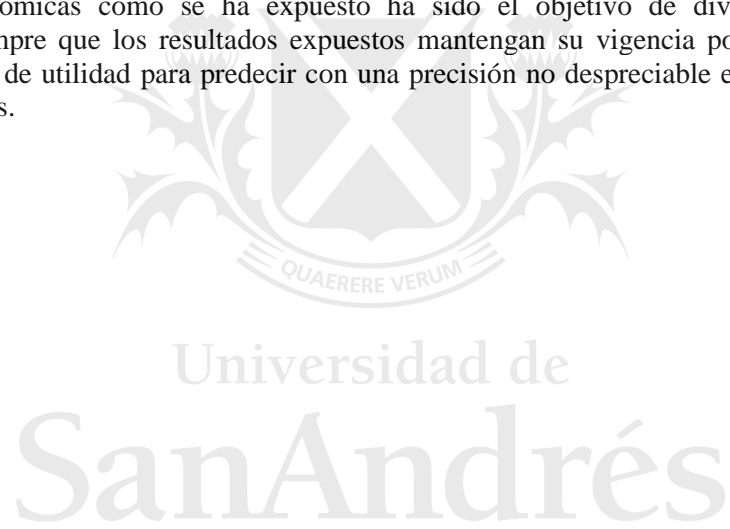
Gráfico 11. Muestra aleatoria de curvas diarias obtenidas por PCA y las correspondientes curvas interpoladas originales



5- Conclusiones

En la presente investigación, se expuso una serie de estudios en los cuales se realizaron PCA para las curvas de rendimientos soberanas de distintos países y momentos históricos y se describió la metodología más popularmente utilizada en los mismos. Dicho procedimiento se realizó en el presente caso con cotizaciones de bonos soberanos en moneda extranjera correspondientes al período 2014-2019, encontrando que el resultado está en línea con la literatura en la materia. El primer componente, o factor “nivel”, explicó el 89,0%, y en conjunto con los componentes vinculados a pendiente y curvatura explicaron el 97,9% de la volatilidad de la curva de rendimientos. Esto indica que una estrategia de cobertura por *duration* sería suficiente para, en el largo plazo, compensar casi el 90% de la volatilidad de la curva, aún a través de períodos como el seleccionado durante el cual pueden señalarse al menos dos episodios puntuales en los que la pendiente de la curva se invirtió.

En investigaciones posteriores, la utilización de los componentes obtenidos puede permitir elaborar modelos de pronóstico que podrán computarse con relativa facilidad, especialmente si se los obtiene computando interpolaciones por la metodología de Nelson y Siegel. Estos modelos pueden ser de utilidad para mensurar la manera en que los mercados perciben y reaccionan ante los cambios en determinadas variables macroeconómicas como se ha expuesto ha sido el objetivo de diversas investigaciones. Adicionalmente, siempre que los resultados expuestos mantengan su vigencia por fuera de la ventana escogida, pueden ser de utilidad para predecir con una precisión no despreciable el sendero de la propia curva de rendimientos.



6- Bibliografía

- Abbritti, Mirko, Salvatore Dell'Erba, Antonio Moreno, and Sergio Sola. 2012. "Global Factors in the Term Structure of Interest Rates."
- Afonso, Antonio, and Manuel Martins. 2010. "Level, slope, curvature of the sovereign yield curve and fiscal behaviour." *European Central Bank*.
- Cortés Espada, Josué F., Manuel Ramos-Francia, and Alberto Torres García. 2008. "An Empirical Analysis of the Mexican Term Structure of Interest Rates." *Documentos de Investigación - Banco de México*.
- Crump, Richard K., and Nikolay Gospodinov. 2019. "Deconstructing the Yield Curve." *Staff Reports - Federal Reserve of New York*.
- Delfiner, Miguel T. 2004. "Patrones de fluctuación de la curva de rendimientos en Argentina."
- Diebold, Francis, and Canlin Li. 2003. "Forecasting the term structure of government bond yields." *National Bureau of Economic Research*.
- Diebold, Francis X., Lei Ji, and Canlin Li. 2004. "A Three-Factor Yield Curve Model: Non-Affine Structure, Systematic Risk Sources, and Generalized Duration."
- Ehrmann, Michael, Fratzscher Marcel, Refet S. Gurkaynak, and Swanson T. Eric. 2007. "Convergence and anchoring of yield curves in the Euro area." *European Central Bank - Working Paper Series*.
- Favero, Carlo A., and Stefano W. Giglio. 2006. "Fiscal Policy and the Term Structure: Evidence from the Case of Italy in the EMS and the EMU Periods." *Working Paper Series - Innocenzo Gasparini Institute for Economic Research*.
- Karelmo, Kia. 2010. "Interest Rate Term Structure Modeling in the Presence of Missing Data."
- Martha, Mota Aragón B., and Mata Mata Leovardo. n.d. "The yield curve and its relation to economic activity in Mexico." *Contaduría y Administración - Universidad Nacional Autónoma de México*.
- Nath, Golaka. 2012. "Estimating Term Structure Changes using Principal Component Analysis in Indian Sovereign Bond Market." *Munich Personal RePEc Archive*.
- Charles R. Nelson and Andrew F. Siegel. 1987. "Parsimonious Modeling of Yield Curves" *The University of Chicago Press*.
- Rebonato, Riccardo. 2018. "Bond Pricing and Yield Curve Modeling." *Cambridge University Press*.
- Scheinkman, José, and Robert Litterman. 1991. "Common factors affecting bond returns." *The Journal of Fixed Income*.
- Suimon, Yoshiyuki, Hiroki Sakaji, Kiyoshi Izumi, and Hiroyasu Matsushima. n.d. "Autoencoder-Based Three-Factor Model for the Yield Curve of Japanese Government Bonds and a Trading Strategy." *Journal of Risk and Financial Management*.

7- Anexo

Se presentan los scripts utilizados en lenguaje Python, para la realización de los cálculos expuestos en las secciones previas.

1) Cálculo de los valores diarios de TIR y MD para cada activo:

```
#Importación de paquetes necesarios

import pandas as pd
import scipy.optimize
from datetime import date
import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')
import numpy as np

"""Importación y formateo de precios"""

precios_path = "/RUTA/"
precios= pd.read_csv(precios_path + "preciosfiltrados.csv",parse_dates=["Fecha"])
precios=precios.iloc[:,1:]
precios=precios[["Fecha", "Especie", "Último"]]

"""Importación de flujos de pagos"""

flujos_path = "/content/drive/My Drive/RUTA/"
flujos= pd.read_excel(flujos_path + "Activos flujos.xlsx",sheet_name="Flujo unico")

"""Definición de funciones XIRR y NPV"""

def xnpv(rate, values, dates):
    if rate <= -1.0:
        return float('inf')
    d0 = dates[0] # or min(dates)
    return sum([ vi / (1.0 + rate)**((di - d0).days / 365.0) for vi, di in zip(values, dates)])
#el commando zip contruye tuplas iterables con ambas listas presentadas

def xirr(values, dates):
    try:
        return scipy.optimize.newton(lambda r: xnpv(r, values, dates), 0.0)
    except RuntimeError: # Failed to converge?
        return scipy.optimize.brentq(lambda r: xnpv(r, values, dates), -1.0, 1e10)

precios.set_index("Fecha",inplace=True)
flujos=flujos[["Fecha","Ticker","PAGO"]]
flujos.set_index(keys=["Fecha","Ticker"],inplace=True)

#Asociamos a cada registro de cotización un subconjunto de flujos correspondiente a los pagos posteriores a esa fecha, para ese bono, y sus fechas correspondientes

precios["TIR"]=0.0
precios.reset_index(inplace=True)
flujos.reset_index(inplace=True)
precios["TIR"].astype(object)
precios=precios[["Fecha","Especie","Último"]]

#Se remueven precios y flujos nulos (observados en los días de emisión de los bonos)
```

```

precios=precios[precios["Último"]!=0]
flujos=flujos[flujos["PAGO"]!=0]

#Se calculan las TIR para cada fecha y cada activo

for i in range(len(precios)):
    fechaflujos=flujos[(flujos["Fecha"]>precios.iloc[i,0])&(flujos["Ticker"]==precios.iloc[i,1])][["Fecha","PAGO"]]
    a=pd.Series(fechaflujos["Fecha"])
    b=pd.Series(fechaflujos["PAGO"])
    a=(a.tolist()+a.iloc[0])
    b=(b.tolist()+b.iloc[0])
    precios.iloc[i,3]=xirr(b,a)

#Para graficar histogramas e identificar outliers:

x = preciosB["TIR"]
plt.hist(x, bins = 10)
plt.show()

preciosB[preciosB["TIR"]>0.2]
preciosB=preciosB[preciosB["TIR"]<0.5]

#Se descartan registros con TIR mayor a 50% (percentil 99%).

x = preciosB["TIR"]
plt.hist(x, bins = 50)
plt.show()

preciosB=preciosB[["Fecha","Especie","Último","TIR"]]

#Cálculo de Modified Duration

from datetime import datetime
preciosB["MD"]=0.0
c=[]
for i in range(len(preciosB)):
    fechaflujos=flujos[(flujos["Fecha"]>preciosB.iloc[i,0])&(flujos["Ticker"]==preciosB.iloc[i,1])][["Fecha","PAGO"]]
    a=pd.Series(fechaflujos["Fecha"])
    b=pd.Series(fechaflujos["PAGO"])

    for j in range(len(a)):
        a.iloc[j]=(a.iloc[j]-preciosB.iloc[i,0]).days/365
        b.iloc[j]=(b.iloc[j]/(1+preciosB.iloc[i,3])**a.iloc[j])
        c.append(a.iloc[j]*b.iloc[j])
    preciosB.iloc[i,4]=sum(c)/preciosB.iloc[i,2]

c.clear()

#Cálculo de TTM

from datetime import datetime
preciosB["TTM"]=0
preciosB["Vencimiento"]=0
preciosB["Vencimiento"]=pd.Series(preciosB["Vencimiento"])
preciosB["Fecha"]=pd.Series(preciosB["Fecha"])

for i in range(len(preciosB)):
    preciosB.iloc[i,6]=flujos[flujos["Ticker"]==preciosB.iloc[i,1]][["Fecha"].max()]

for i in range(len(preciosB)):
    preciosB.iloc[i,6]=datetime.strptime(str(preciosB.iloc[i,6]):10,"%Y-%m-%d")

```

```

for i in range(len(preciosB)):
    preciosB.iloc[i,0]=datetime.strptime(str(preciosB.iloc[i,0]):10,"%Y-%m-%d")

for i in range(len(preciosB)):
    preciosB.iloc[i,0]=datetime.strptime(str(preciosB.iloc[i,0]),"%Y-%m-%d %H:%M:%S")
    preciosB.iloc[i,6]=datetime.strptime(str(preciosB.iloc[i,6]),"%Y-%m-%d %H:%M:%S")

pd.Series(preciosB["Vencimiento"])
pd.Series(preciosB["Fecha"])
preciosB["Vencimiento"]=pd.to_datetime(preciosB["Vencimiento"])
preciosB["TTM"]=preciosB["Vencimiento"]-preciosB["Fecha"]
for i in range(len(preciosB)):
    preciosB.iloc[i,5]=(preciosB.iloc[i,6]-preciosB.iloc[i,0]).days

#Exportación

preciosB["TTM"] = pd.to_numeric(preciosB["TTM"], downcast="float")
preciosB.to_csv(r" /RUTA/BaseTIR_MD.csv")

```

2) Interpolación de curvas por Cubic Splines:

```

#Instalación a importación de paquetes necesarios.
!pip install nelson_siegel_svansson
!pip install gekko
!pip install --upgrade gekko

from gekko import GEKKO
import numpy as np
import matplotlib.pyplot as plt
from datetime import datetime
now = datetime.now()
import pylab as plb
import pandas as pd

#Importación de base TIR-MD
TIRes_path = "/RUTA/"
tiresmdcompleto= pd.read_csv(TIRes_path + "BaseTIR_MD.csv")
tiresmdcompleto= tiresmdcompleto.loc[:,["Fecha","Especie","TIR","TTM","MD"]]

tiresmdcompleto.head()

#Se acortan las fechas, quitando horas, minutos y segundos

for i in range(len(tiresmdcompleto)):
    tiresmdcompleto.iloc[i,0]=tiresmdcompleto.iloc[i,0]:10]
    print(tiresmdcompleto["Fecha"].max()+ " "+tiresmdcompleto["Fecha"].min())

#Se definen los extremos de la interpolación

tiresmdcompleto[tiresmdcompleto["Especie"]=="PARAD"].MD.min()
tiresmdcompleto[tiresmdcompleto["Especie"]=="AN18D"].MD.max()

#Se limita la ventana temporal para evitar extrapolación

tiresmdcompleto[tiresmdcompleto["Fecha"]=="2015-12-02"]

#Se clasifican los activos en Ley Argentina y Ley Extranjera

```

```

ext=["AA26D", "AA46D", "A2E2D", "A2E3D", "A2E7D", "A2E8D", "AE48D", "AC17D"]
arg=["AA17D", "AN18D", "AO20D", "AY24D", "PARAD", "DICAD", "AA25D", "AA37D"]
cartera=["AA17D", "AN18D", "AO20D", "AY24D", "PARAD", "DICAD",
"AA37D", "A2E2D", "A2E2D", "A2E3D", "A2E8D", "AE48D", "AC17D", "A2E7D", "AA25D"]

```

#Se construyen los dos sets separados:

```

tiresmdext=tiresmdcompleto[tiresmdcompleto["Especie"].isin(ext)]
tiresmdarg=tiresmdcompleto[tiresmdcompleto["Especie"].isin(arg)]

```

```

tiresmdarg=tiresmdarg[tiresmdarg["TIR"]<=0.17] # Se eliminan valores extremos

```

```

filtroarg=tiresmdarg.groupby(["Fecha"]).count()
fechasarg=list(filtroarg[filtroarg["Especie"]>=6].index)

```

```

filtroext=tiresmdext.groupby(["Fecha"]).count()
fechasext=list(filtroext[filtroext["Especie"]>=6].index)

```

#Comparación Ley Arg, Ley Ext.

```

fechacompara="2018-10-30"

```

#Cubic Spline para Ley Arg

```

xm = tiresmdarg[tiresmdarg["Fecha"]==fechacompara]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==fechacompara]["TIR"].to_numpy()

```

```

m = GEKKO()
m.x = m.Param(value=np.linspace(1,11,11)) #Espacio lineal sobre el que se calculará la función resultante
m.y = m.Var() #Predicciones
m.options.IMODE=2 #Modo "Solución"
m.cspline(m.x,m.y,xm,ym) #Cubic Spline
m.solve(disp=False) #Solve
#help(m.cspline)

```

#Cubic Spline para Ley Ext

```

xmext = tiresmdext[tiresmdext["Fecha"]==fechacompara]["MD"].to_numpy()
ymext = tiresmdext[tiresmdext["Fecha"]==fechacompara]["TIR"].to_numpy()
mext = GEKKO() #CREA MODELO GEKKO
mext.x = mext.Param(value=np.linspace(1,11,11))
mext.y = mext.Var()
mext.options.IMODE=2
mext.cspline(mext.x,mext.y,xmext,ymext)
mext.solve(disp=False)

```

#Gráfico comparativo

```

plt.plot(xmext,ymext,'ko',label='data ext')
plt.plot(mext.x.value,mext.y.value,'y--',label='cubic spline ext')
plt.plot(xm,ym,'bo',label='data arg')
plt.plot(m.x.value,m.y.value,'r--',label='cubic spline arg')
plt.xlabel("MD",size=12)
plt.ylabel("TIR",size=12)
plt.legend()
plt.show()

```

Interpolación para cada día de la ventana temporal escogida, en base a activos Ley Argentina.

```

current_time = now.strftime("%H:%M:%S")
print("Current Time =", current_time)
yt=[]
yt.clear()
for i in fechasarg:

```

```

dia=i
m = GEKKO()
m.x = m.Param(value=np.linspace(0,12,49))
m.y = m.Var()
m.options.IMODE=2
xm = tiresmdarg[tiresmdarg["Fecha"]==i]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==i]["TIR"].to_numpy()

m.cspline(m.x,m.y,xm,ym)
m.solve(dis=False)

yt.append(list((dia,m.x,m.y)))

```

#Se guardan los resultados en un Dataframe y se exportan.

```

datos=pd.DataFrame(yt)
datos.to_csv(TIRes_path+"SETs.csv")

```

#Para graficar ejemplos aleatorios de curvas diarias

```

import random
fig, axs = plt.subplots(5,2,sharex=True,sharey=True)
axs.setx
ies=np.random.randint(1,len(yt),size=10)

i=ies[0]
axs[0,0].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[0,0].plot(xm,ym,'bo',label='data')
axs[0,0].legend()

i=ies[1]
axs[1,0].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[1,0].plot(xm,ym,'bo',label='data')
axs[1,0].legend()

i=ies[2]
axs[2,0].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[2,0].plot(xm,ym,'bo',label='data')
axs[2,0].legend()

i=ies[3]
axs[3,0].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[3,0].plot(xm,ym,'bo',label='data')
axs[3,0].legend()

i=ies[4]
axs[4,0].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tiresmdarg[tiresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[4,0].plot(xm,ym,'bo',label='data')
axs[4,0].legend()

```

```

i=ies[5]
axs[0,1].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[0,1].plot(xm,ym,'bo',label='data')
axs[0,1].legend()

```

```

i=ies[6]
axs[1,1].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[1,1].plot(xm,ym,'bo',label='data')
axs[1,1].legend()

```

```

i=ies[7]
axs[2,1].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[2,1].plot(xm,ym,'bo',label='data')
axs[2,1].legend()

```

```

i=ies[8]
axs[3,1].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[3,1].plot(xm,ym,'bo',label='data')
axs[3,1].legend()

```

```

i=ies[9]
axs[4,1].plot(yt[0][1],pd.to_numeric(yt[i][2]),'r--',label=yt[i][0])
xm = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["MD"].to_numpy()
ym = tioresmdarg[tioresmdarg["Fecha"]==yt[i][0]]["TIR"].to_numpy()
axs[4,1].plot(xm,ym,'bo',label='data')
axs[4,1].legend()

```

```

fig.set_figheight(15)
fig.set_figwidth(15)
plt.legend()

```

```
plt.show()
```

```
# Construcción de nuevos Dataframes y exportación
```

```

final=yt
cuadro=[]
cuadro.clear()
for i in range(len(final)):
    cuadro.append([final[i][0]]+[final[i][2]])

```

```

a=[]
a.clear()
for i in range(len(cuadro)):
    b=[]
    b.clear()
    b.append(cuadro[i][0])
    for j in range(len(cuadro[i][1])):
        b.append(cuadro[i][1][j])
    a.append(b)

```

```

df = pd.DataFrame.from_records(a)
df

```




```

titulos=list(yt[0][1])
titulosbis=["Fecha"]+titulos

df.columns=titulosbis
pd.to_numeric(df.iloc[2,1:])

""EXPORTACION""

Compendio=df
CompendioTIR=Compendio.copy()
CompendioTIR.to_csv(TIRes_path+"TIRes.csv")

```

3) Estimación de PCA

```

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import pylab as plb
from pylab import rcParams
from matplotlib.ticker import (MultipleLocator, AutoMinorLocator)

#Importación de base TIR
path = "/RUTA/"
tirs= pd.read_csv(path + "TIRes.csv")

#Gráfico de evolución de las tasas correspondientes a 2, 5 y 10 años

fig, ax = plt.subplots()
ax.xaxis.set_major_locator(MultipleLocator(60))
ax.plot(tirs["Fecha"],tirs["2.0"],label="2 años")
ax.plot(tirs["Fecha"],tirs["5.0"],label="5 años")
ax.plot(tirs["Fecha"],tirs["10.0"],label="10 años")
ax.set_xlabel("MD",size=12)
ax.set_ylabel("TIR",size=12)
ax.legend()
plt.xticks(rotation=90)
plt.show()

# Implementación del PCA

from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA

x=tirs.iloc[:,2:] #Se remueven las columnas índice y fechas

#Se estandariza el dataset, restando media y dividiendo por desvío estándar.

x = StandardScaler().fit_transform(x)
pcatir = PCA(n_components=3) #Elección del número de componentes de interés, según la varianza explicada incremental

principalComponentsTIR = pcatir.fit_transform(x)

#Se estima el modelo y se alojan en una variable los resultados

#Se alojan los resultados, valores de los componentes para cada fecha de la ventana, en un dataframe.

principalDftir = pd.DataFrame(data = principalComponentsTIR
, columns = ['PC1', 'PC2', 'PC3'])

```

```
#Se obtienen los resultados de interés
```

```
print(pcatir.singular_values_)
pcatir.explained_variance_
PCs=pd.DataFrame(pcatir.components_)
pcatir.explained_variance_ratio_
pcatir.explained_variance_ratio_[0]*100+pcatir.explained_variance_ratio_[1]*100+pcatir.explained_variance_ratio_[2]*100
pcatir.explained_variance_ratio_[0]*100+pcatir.explained_variance_ratio_[1]*100
```

```
#Se grafican los resultados
```

```
rcParams['figure.figsize'] = 10, 5
t = tirs.Fecha
plt.plot(t,principalDftir["PC1"],"b--",label="PC1")
plt.plot(t,principalDftir["PC2"],"r*",label="PC2")
plt.plot(t,principalDftir["PC3"],"y*",label="PC3")
plt.legend()
plt.xticks(np.arange(0,len(vars),60),rotation="vertical")
plt.show()
```

```
fig, axs = plt.subplots(3,1,sharex=True)
t = tirs.Fecha
```

```
axs[0].plot(t,principalDftir["PC1"],"b--",label="PC1")
axs[0].legend(loc=3)
```

```
axs[1].plot(t,principalDftir["PC2"],"r*",label="PC2")
axs[1].legend(loc=3)
```

```
axs[2].plot(t,principalDftir["PC3"],"yo",label="PC3")
axs[2].legend(loc=3)
```

```
fig.set_figheight(15)
fig.set_figwidth(15)
plt.xticks(np.arange(0,len(tirs),60),rotation="vertical")
plt.show()
```

