



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**“DESARROLLO DE UNA METODOLOGÍA DE ASSET ALLOCATION PARA  
UNA AFP”**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

CARLOS BASTIÁN LEA GRENETT

PROFESOR GUÍA:  
JAVIER SUAZO SÁEZ

MIEMBROS DE LA COMISIÓN:  
CARLOS PULGAR ARATA  
GUILLERMO ESPINOZA ARMIJO

SANTIAGO DE CHILE  
2020

## **RESUMEN DE LA MEMORIA PARA OPTAR AL**

**TÍTULO DE:** Ingeniero Civil Industrial

**POR:** Carlos Bastián Lea Grenett

**FECHA:** 13/01/2020

**PROFESOR GUÍA:** Javier Suazo Sáez

### **DESARROLLO DE UNA METODOLOGÍA DE ASSET ALLOCATION PARA UNA AFP**

Actualmente el país se encuentra en una crisis social en la que las administradoras de fondos de pensiones están fuertemente cuestionadas, debido principalmente a su desempeño y las jubilaciones que reciben sus afiliados. Con esto en consideración, es posible evidenciar que gran parte del éxito de estas depende de que tan bien manejan sus inversiones y no de la proporción del mercado de activos que poseen, por lo que existe un valor agregado en innovar e ir a la vanguardia en lo que respecta a inversión, más aún, si se considera que aumentar en 10 puntos base la rentabilidad sobre el patrimonio inicial de la compañía significan alrededor de 767 millones de pesos extras en sus resultados anuales.

La metodología utilizada en el trabajo de título es propia de la ciencia de los datos, la cual se identifica como un proceso iterativo en el que se está estudiando, desarrollando, testeando y mejorando continuamente. Se tiene como resultado una recomendación de inversión a nivel top del proceso de asset allocation, es decir, como repartir el total de fondos asociado a cada tipo de multifondo en las distintas clases de activos en los que la administradora invierte el dinero de los afiliados, los cuales se identifican como: renta fija local, renta fija investment grade, renta fija high yield, renta variable de mercados emergentes y renta variable de mercados desarrollados.

El Framework de trabajo se basa en un desarrollo realizado en el año 2017, implicando que se esté a la vanguardia del asset allocation. Dicho esto, se desarrollan distintas estrategias que capturan diferentes características que afectan a los activos financieros. La principal cualidad de las estrategias desarrolladas se basa en que una captura el crecimiento económico utilizando un horizonte de tiempo que va desde los 6 meses y, la otra, captura la tendencia en el mercado utilizando un horizonte de tiempo que va desde los 8 días.

Se destaca la obtención de estrategias ortogonales entre sí, lo que se evidencia a través de correlaciones cercanas a 0 encontradas, resultando que la mezcla de estas es capaz de obtener, en general, un performance positivo consistente y persistente. Este hallazgo viene a ser un una extensión de lo que se entiende comúnmente como diversificación de riesgo, puesto que junto con diversificar invirtiendo en distintos tipos de activos, se diversifica invirtiendo en distintos tipos de estrategias, obteniéndose resultados más robustos.

## **AGRADECIMIENTOS**

En primer lugar, quienes me hacen sentir afortunado por tenerlos a mi lado y me llenan de cariño desde el día 0; Mamá, hermana, abuelos, tíos, primos y sobrino.

En segundo lugar, quienes me hacen feliz y me entregan amor; Rocío y Lincoln.

En tercer lugar, quienes me acompañan de otra forma; Benito, Inara y Belén.

En cuarto lugar, los de siempre; memes (malos) y videos, old school, piscoleros y bullshef.

En quinto lugar, quienes de distinta manera hicieron el proyecto posible; Guillermo, Carlos, Javier, Víctor y Karla.

En sexto lugar, quienes se ganaron una parte de mí; Ure y Rosita.

No puedo terminar este trabajo sin darle un espacio a ustedes que, a su manera, tuvieron una labor positiva en lo que fue mi trabajo de título, formación profesional y mi vida en general.

## TABLA DE CONTENIDO

<b>1. INTRODUCCIÓN</b> .....	<b>1</b>
1.1. Industria de la AFP .....	1
1.2. Situación Actual del Sistema de Pensiones.....	4
1.3. Metodologías Cuantitativas en Inversión .....	5
1.4. Crisis Social en Chile: ¿Cómo Afecta a la Problemática? .....	7
<b>2. PROYECTO Y JUSTIFICACIÓN</b> .....	<b>8</b>
2.1. Oportunidad de Mejora.....	8
2.2. Hipótesis y Alternativas de Solución.....	9
<b>3. OBJETIVOS</b> .....	<b>11</b>
3.1. Objetivo General .....	11
3.2. Objetivos Específicos.....	11
<b>4. ALCANCES</b> .....	<b>12</b>
<b>5. RESULTADO DE LA METODOLOGÍA</b> .....	<b>13</b>
<b>6. MARCO TEÓRICO</b> .....	<b>14</b>
6.1. Framework.....	14
6.2. Optimización de Portafolios .....	15
6.3. Estrategia Económica .....	16
6.3.1. Clustering: KM y GMM.....	16
6.4. Estrategia Momentum .....	18
6.4.1. MACD .....	18
<b>7. METODOLOGÍA</b> .....	<b>20</b>
<b>8. DESARROLLO DEL PROYECTO</b> .....	<b>22</b>
8.1. Índices.....	22
8.2. Modelo Económico .....	23
8.2.1. Data .....	23
8.2.2. Tratamiento de Datos .....	25
8.2.3. Clustering de Fase y Probabilidad de Pertenencia.....	30
8.2.4. Cálculo y Optimización del Portafolio final.....	34
8.2.5. Análisis y Backtest.....	41
8.2.6. K-means .....	43
8.2.7. Análisis de sensibilidad – KM .....	50
8.2.8. Síntesis - KM .....	59
8.2.9. GMM .....	61
8.2.10. Análisis de Sensibilidad - GMM .....	69
8.2.11. Síntesis - GMM .....	75
8.3. Modelo Momentum .....	76

8.3.1.	Data .....	76
8.3.2.	Tratamiento de Datos .....	76
8.3.3.	Señal – Tendencia .....	78
8.3.4.	Construcción de Portafolio.....	80
8.3.5.	Análisis y Backtest.....	81
8.3.6.	Análisis de Sensibilidad .....	84
8.3.7.	Síntesis - Momentum .....	91
<b>8.4.</b>	<b>Mezcla de Modelos .....</b>	<b>92</b>
8.4.1.	Correlación .....	94
8.4.2.	Performance y Métricas.....	95
8.4.3.	Síntesis – Mezcla de Modelos .....	100
<b>9.</b>	<b>CONCLUSIÓN .....</b>	<b>101</b>
<b>10.</b>	<b>BIBLIOGRAFÍA.....</b>	<b>103</b>
<b>11.</b>	<b>ANEXOS .....</b>	<b>104</b>
11.1.	Anexo 1 .....	104
11.2.	Anexo 2 .....	106
11.3.	Anexo 3 .....	106
11.4.	Anexo 4 .....	107
11.5.	Anexo 5 .....	109
11.6.	Anexo 6 .....	110

## 1. INTRODUCCIÓN

Los fondos del sistema de capitalización individual obligatoria son manejados por administradoras de fondos de pensión (en adelante AFP), quienes según normas y restricciones impuestas por la superintendencia de pensiones<sup>1</sup> (en adelante SP) manejan el dinero de sus afiliados con el fin de asegurar una pensión al momento de jubilar. Por lo tanto, es relevante la performance que tenga la AFP con respecto a las inversiones que realiza para aumentar el capital del afiliado, impactando el bienestar social del país.

Con esto en consideración, resulta relevante para una AFP poder diferenciarse del resto, para así obtener mayor cantidad de clientes y contribuir de mejor manera en el bienestar de los pensionados. Una característica diferenciadora es la performance de las inversiones que realiza AFP, con esto en consideración, resulta vital el tomar decisiones informadas y certeras tal que la utilidad de la compañía sea mayor, a su vez, estas decisiones pueden ser basadas en juicios de expertos, conocimiento de rubros, modelos matemáticos/estadísticos, etc. Donde el desarrollo del trabajo de título se centra en este último criterio.

A su vez, el mundo financiero es sumamente volátil, por lo que es complejo generar alguna técnica de cualquier tipo que sea capaz de generar rentabilidad en todo periodo de tiempo, por lo tanto, resulta sumamente interesante la idea de encontrar técnicas que sean complementarias entre sí, es decir, que cuando una no genera retorno positivo la otra sí y viceversa, de manera de obtener una mezcla tal que en el largo plazo posea pendiente positiva.

Por otro lado, es relevante destacar que, aunque el proyecto es desarrollado para una AFP, es funcional y útil para cualquier compañía de inversiones a nivel nacional e internacional, puesto que está diseñado para tomar mejores decisiones de inversión a partir de la información existente.

### 1.1. Industria de la AFP

La industria de las AFPs está conformada por: AFP Capital, AFP Cuprum, AFP Habitat, AFP Modelo, AFP Planvital y AFP Provida.

Según las restricciones y regulaciones impuestas por la SP se establecen los tipos de fondos en los que cada afiliado puede invertir, y que constituyen el eje de acción principal de las AFPs. Cada fondo se diferencia por el riesgo que quiera asumir el afiliado. La siguiente figura muestra características de los fondos:

---

<sup>1</sup> Superintendencia de pensiones. Institución. Disponible en: <http://www.spensiones.cl/portal/institucional/594/w3-propertyvalue-5990.html> (consulta: 1 de mayo de 2019).

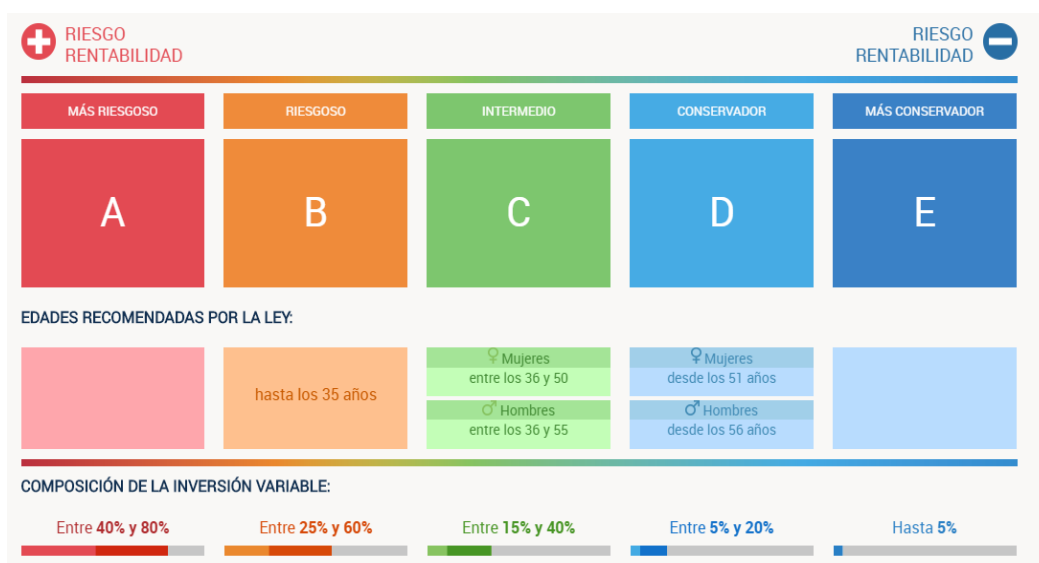


Figura 1: Tipos de Fondos<sup>2</sup>.

La siguiente tabla resumen muestra que proporción del mercado se reparte cada AFP del total de afiliados a diciembre de 2018, que ascienden a la suma de 10.705.051. Donde se identifica a AFP Provida como la más grande con el 28% del total de afiliados y AFP Cuprum como la más pequeña con un 6% del total:

	<b>Afiliados</b>	<b>Porcentaje</b>
<b>AFP Provida</b>	2.997.462	28%
<b>AFP Habitat</b>	1.971.368	18%
<b>AFP Plan Vital</b>	1.775.737	17%
<b>AFP Modelo</b>	1.730.618	16%
<b>AFP Capital</b>	1.629.979	15%
<b>AFP Cuprum</b>	599.885	6%

Tabla 1: Afiliados por AFP<sup>3</sup>.

Sin embargo, si se analiza la cantidad de activos financieros que maneja cada AFP a la misma fecha, se destaca que AFP Cuprum maneja el 19% del mercado en ahorro obligatorio, gestionando \$25.501.919 millones de pesos<sup>4</sup>. En la siguiente tabla se adjunta la participación de mercado de ahorro obligatorio, donde se destacan AFP Habitat y AFP Provida.

<sup>2</sup> AFP Habitat. Multifondos. Disponible en: <https://www.afphabitat.cl/multifondos/> (consulta: 1 de mayo de 2019).

<sup>3</sup> Superintendencia de pensiones. Afiliados AFP. Disponible en: [https://www.spensiones.cl/apps/loadEstadisticas/genEstadAfiliadosCotizantes.php?id=inf\\_estadistica/aficot/mensual/2018/12/01C.html&p=M&menu=sci&menuN1=afil&menuN2=afp&orden=10&ext=.html](https://www.spensiones.cl/apps/loadEstadisticas/genEstadAfiliadosCotizantes.php?id=inf_estadistica/aficot/mensual/2018/12/01C.html&p=M&menu=sci&menuN1=afil&menuN2=afp&orden=10&ext=.html) (consulta: 1 de mayo de 2019).

<sup>4</sup> AFP Cuprum. Memoria 2018. Disponible en: [https://nuevo.cuprum.cl/sites/default/files/Memoria%20Cuprum%202018\\_1.pdf](https://nuevo.cuprum.cl/sites/default/files/Memoria%20Cuprum%202018_1.pdf) (consulta: 1 de mayo de 2019).

	<b>Dinero en Activos [Millones CLP]</b>	<b>Tamaño de Mercado</b>
<b>AFP Provida</b>	\$ 33.474.127	25%
<b>AFP Habitat</b>	\$ 37.594.407	28%
<b>AFP Plan Vital</b>	\$ 4.556.580	3%
<b>AFP Modelo</b>	\$ 7.489.877	6%
<b>AFP Capital</b>	\$ 25.727.805	19%
<b>AFP Cuprum</b>	\$ 25.501.919	19%

Tabla 2: Cantidad de Activos por AFP<sup>5</sup>.

Por último, es relevante mencionar la rentabilidad anual sobre el patrimonio neto inicial de cada AFP, para así ajustar los resultados al tamaño de cada compañía. Se destaca que AFP Cuprum es una de las con menor rentabilidad, superando solo a AFP Provida y muy lejos a AFP Habitat, quien maneja la mayor proporción de activos del mercado:

	<b>AFP Capital</b>	<b>AFP Cuprum</b>	<b>AFP Habitat</b>	<b>AFP Modelo</b>	<b>AFP Planvital</b>	<b>AFP Provida</b>
<b>2018</b>	11,89%	6,04%	25,14%	60,31%	26,10%	5,54%
<b>2017</b>	15,25%	7,90%	32,82%	94,42%	6,19%	6,93%

Tabla 3: Rentabilidad de las AFP Según Patrimonio Inicial<sup>6</sup>.

Por lo tanto, es evidente que la performance de cada compañía depende de que tan bien se tomen las decisiones de inversión. Con esto en consideración, existe el constante desafío de innovar y mejorar los procesos actuales de toma de decisión, donde las metodologías cuantitativas están tomando un rol relevante.

<sup>5</sup> Superintendencia de pensiones. Activos AFP. Disponible en: <https://www.spensiones.cl/apps/centroEstadisticas/paginaCuadrosCCEE.php?menu=sest&menuN1=sistpens&menuN2=fondospem> (consulta: 2 de mayo de 2019).

<sup>6</sup> Superintendencia de pensiones. Rentabilidad AFP. Disponible en: [https://www.spensiones.cl/apps/loadEstadisticas/loadEstadFinancieras.php?id=inf\\_estadistica/iftafp/2018/12/rentab201812.html&menu=sci&menuN1=estfinafp&menuN2=NOID&orden=10&ext=.html](https://www.spensiones.cl/apps/loadEstadisticas/loadEstadFinancieras.php?id=inf_estadistica/iftafp/2018/12/rentab201812.html&menu=sci&menuN1=estfinafp&menuN2=NOID&orden=10&ext=.html) (consulta: 2 de mayo de 2019).



## 1.2. Situación Actual del Sistema de Pensiones

Para empezar, es importante tener en claro los pilares del sistema de pensiones de Chile, los que actualmente son<sup>7</sup>:

- Pilar Contributivo Obligatorio: Basado en la capitalización del ahorro obligatorio de cada trabajador.
- Pilar Contributivo Voluntario: Basado en la capitalización del ahorro complementario al obligatorio de cada trabajador.
- Pilar Solidario: Dirigido al 60% más pobre del país. Es financiado por el estado.

Bajo esta lógica, el sistema actual de pensiones se encuentra cuestionado, principalmente por los bajos montos que reciben los afiliados a este. Según diario Uchile, el 60% de los adultos mayores recibe una jubilación mensual entre \$50.000 y \$150.000 CLP, más aún, un 59% de las mujeres recibe pensiones bajo la línea de la pobreza<sup>8</sup>. Cifras que evidencian que algo no funciona correctamente.

A su vez, se expone la problemática de la regularización de las AFP dado el poder político y económico que adquieren al manejar fondos que superan el PIB de Chile y, además, ser controladas, en general, por multinacionales extranjeras (Provida, Capital, Cuprum y Plan Vital)<sup>9</sup>.

Lo anterior ha implicado una crisis, provocando la manifestación de los ciudadanos y un descontento generalizado<sup>10 11</sup>. Es así como surgió el movimiento “No + AFP” que propone como solución un sistema de pensiones de reparto solidario, tripartito y administrado por el estado, es decir, un sistema en el que el estado, el empleador y el contribuyente impongan en un fondo común del que se financiaría cada jubilación<sup>12 13</sup>.

---

<sup>7</sup> Previsión Social. Sistema de Pensiones. Disponible en: <https://www.previsionsocial.gob.cl/sps/seguridad-social/sistema-de-pensiones/2/> (consulta: 14 de octubre de 2019).

<sup>8</sup> Diario Uchile. El desolado rostro del sistema de pensiones de Chile. Diponible en: <https://radio.uchile.cl/2018/11/09/el-desolado-rostro-del-sistema-de-pensiones-chileno/> (consulta: 14 octubre de 2019)

<sup>9</sup> Ciper. La Crisis del Sistema Privado de Pensiones, un punto de inflexión de nuestra democracia. Dispobible en: <https://ciperchile.cl/2018/10/30/la-crisis-del-sistema-privado-de-pensiones-un-punto-de-inflexion-de-nuestra-democracia/>

<sup>10</sup> T13. Miles de Personas Marchan Exigiendo Fin AFP. Disponible en: <https://www.t13.cl/noticia/nacional/video-miles-personas-marchan-exigiendo-fin-afp-santiago-y-regiones> (consulta: 14 de octubre de 2019).

<sup>11</sup> El Mostrador. No + AFP cifra en 2.000 asistentes a manifestaciones. Disponible en: <https://www.elmostrador.cl/dia/2019/03/31/noafp-cifra-en-2-mil-los-asistentes-a-manifestacion-contra-el-sistema-de-pensiones-en-santiago/> (consulta: 14 de octubre de 2019).

<sup>12</sup> Coordinadora no más AFP. Inicio. Disponible en: <https://coordinadoranomasafp.cl/> (consulta: 14 de octubre de 2019).

<sup>13</sup> No + AFP. Inicio. Disponible en: <http://www.nomasafp.cl/inicio/>. (consulta: 14 de octubre de 2019).

Por otro lado, se tiene una visión con respecto al problema de las pensiones que se relaciona con la estructura del sistema y no con las administradoras de los fondos, de los cuales se destacan<sup>14</sup>:

- Aumento de la expectativa de vida al momento de jubilar, que se diferencia con el que se prevía al momento de crear el sistema en 1981.
- Caída en las tasas de interés de Chile y el mundo.
- Baja densidad de contribuciones, es decir, un mercado laboral con recurrentes fluctuaciones en la tasa de desempleo e inactividad, sueldos que no presentan imponibles, etc.

Por lo tanto, por un lado es posible visibilizar que el sistema actual no funciona de manera óptima y, por otro lado, que causas de los problemas existentes provienen de características del mercado y de factores que no se pudieron prever de buena manera. Dicho esto y, sea cualquiera el caso, invertir de manera óptima es un aporte en cualquier escenario, lo que es el tema principal de este trabajo de título.

### **1.3. Metodologías Cuantitativas en Inversión**

Según Investopedia, las estrategias cuantitativas han estado presente hace más de 80 años en el mundo de las inversiones mediante modelos matemáticos que detectan oportunidades de inversión. A su vez, estas estrategias pueden fallar, puesto que se basan en eventos pasados y, en general, no consideran eventos futuros posibles, implicando que cuando fallan se obtengan grandes consecuencias<sup>15</sup>.

También, la gran cantidad de datos que día a día se almacenan y el poder de computo de la tecnología actual motiva el uso de estas estrategias con el fin de eliminar la “cuestión sentimental” del proceso de inversión. Por esto es una disciplina efectiva, ya que permite ejecutar decisiones de manera consistente, además, es rentable pudiendo implicar reducción de costos al necesitar una menor cantidad de analistas y/o disminuir la necesidad de viajar y conocer compañías/managers para invertir<sup>16</sup>.

---

<sup>14</sup> CEP Chile. Otra mirada al problema de las pensiones. Disponible en: <https://www.cepchile.cl/otra-mirada-al-problema-de-las-pensiones/cep/2017-08-16/150014.html> (consulta: 14 de octubre de 2019).

<sup>15</sup> Investopedia. Quant Strategies. Disponible en: <https://www.investopedia.com/articles/trading/09/quant-strategies.asp>. (consulta: 14 de octubre de 2019).

<sup>16</sup> Investopedia. Simple Overview Quantitative Analysis. Disponible en: <https://www.investopedia.com/articles/investing/041114/simple-overview-quantitative-analysis.asp> (consulta: 14 de octubre de 2019).

A su vez, existen fondos puramente cuantitativos, los cuales en base a algoritmos sofisticados toman cada una de las decisiones de inversión. El auge de estos viene dado de, tal y como se menciona en el párrafo anterior, la gran cantidad de datos existentes hoy en día, implicando que más compañías se interesen en aumentar su inversión en este tipo de fondos<sup>17</sup>. También, estos fondos son clasificados como inversión alternativa<sup>18</sup>.

Por lo tanto, dado que los métodos cuantitativos no son infalibles y, a su vez, aportan en la generación de exceso de retorno, es razonable que el enfoque que se le den a estas estrategias tenga componentes fundamentales, convergiendo a un proceso “Quantamental”, tal como se muestra en la siguiente figura:

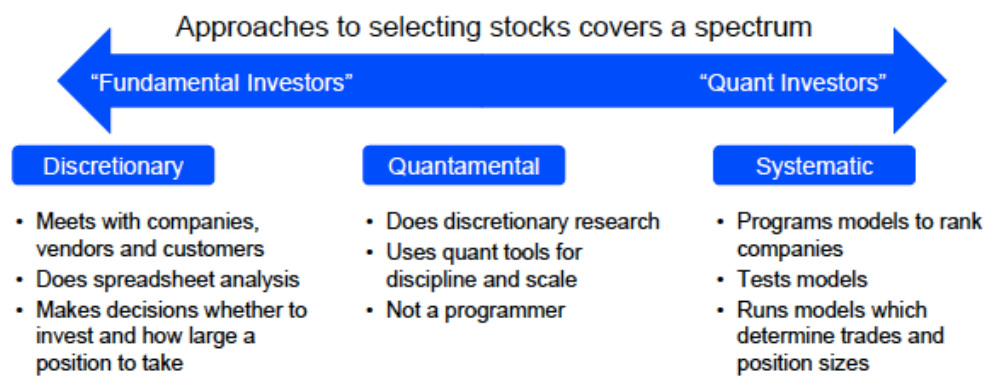


Figura 2: Quantamental<sup>19</sup>.

<sup>17</sup> Pension & Investment. Blackrock Reorganizes Active Equity Unit, Quant Strategies. Disponible en: <https://www.pionline.com/article/20170329/ONLINE/170329859/blackrock-reorganizes-active-equity-unit-turns-more-to-computer-based-quant-strategies> (consulta: 14 de octubre de 2019).

<sup>18</sup> Investopia. Quant Fund. Disponible en: <https://www.investopedia.com/terms/q/quantfund.asp>. (consulta: 14 de octubre de 2019).

<sup>19</sup> Goldman, R. Understanding The Equity Investment Process. Refinitiv, 2019.

#### 1.4. Crisis Social en Chile: ¿Cómo Afecta a la Problemática?

El 6 de octubre de 2019 se anunció un alza de \$30 CLP en el transporte público en la capital de Chile, provocando la organización de distintos grupos, en especial de estudiantes secundarios, para realizar jornadas de evasión masiva en el Metro de Santiago. En tanto, el 18 de octubre de 2019 ante la nula respuesta a las manifestaciones, sumado a desafortunadas declaraciones por parte de, por ejemplo, el ex presidente de Metro, Clemente Pérez, que expuso: “Cabros, esto no prendió”<sup>20</sup>, se desarrolló una de las jornadas de manifestaciones más violenta ocurridas hasta la fecha, implicando para Metro de Santiago un saldo de 9 estaciones incendiadas y 68 con daños<sup>21</sup>.

Luego de los acontecimientos sucedidos ese día, se desencadenó un estallido social a nivel nacional. El estallido tomó la consigna “no son 30 pesos, son 30 años”, haciendo alusión a que el alza \$30 CLP en la tarifa del transporte fue la gota que rebalsó el vaso en un sistema de abusos permanente existentes en el país<sup>22</sup>. A su vez, la ciudadanía expuso las principales razones del descontento que se materializaron en demandas<sup>23</sup>, las que se exponen a continuación:

- Sistema de pensiones.
- Salud y desprotección.
- Transporte público.
- Privatización del agua.
- Educación y movilidad social.
- Abusos y corrupción.
- Salarios.

---

<sup>20</sup> El Mostrador. Cabros Esto no Prendió. Disponible en :<https://www.elmostrador.cl/noticias/multimedia/2019/10/25/cabros-esto-no-prendio-el-dia-en-que-el-expresidente-de-metro-le-bajo-el-perfil-a-las-manifestaciones/> (consulta: 5 de enero de 2020).

<sup>21</sup> Emol. Resumen Crisis Social. Disponible en: <https://www.emol.com/noticias/Nacional/2019/11/18/967500/10-hechos-resumen-crisis-social.html> (consulta: 5 de enero de 2020).

<sup>22</sup> Sputniknews. No Son 30 Pesos Son 30 Años. Disponible en: <https://mundo.sputniknews.com/america-latina/201910221089057254-no-son-30-pesos-son-30-anos-de-abuso-sindicalista-explica-las-protestas-en-chile/> (consulta: 5 de enero de 2020).

<sup>23</sup> T13. Las 6 Grandes Deudas Sociales. Disponible en: <https://www.t13.cl/noticia/nacional/bbc/protestas-en-chile-las-6-grandes-deudas-sociales-por-las-que-muchos-chilenos-dicen-sentirse-abusados> (consulta: 5 de enero de 2020).

De estas, el trabajo de título se corresponde con la demanda relacionada al sistema de pensiones. En esta se expone, tal y como se aprecia en el apartado 1.2., que las pensiones son indignas y que se busca poner fin al sistema de capitalización actual<sup>24</sup>. Con esto en consideración, el trabajo de título se hace cargo, en cierta medida, de la problemática, puesto que no busca modificar el sistema actual que presenta falencias, pero sí busca tomar mejores decisiones de inversión, lo que puede significar aumentar las pensiones actuales o, en su defecto, no hacerlas disminuir considerablemente.

## 2. PROYECTO Y JUSTIFICACIÓN

### 2.1. Oportunidad de Mejora

El proceso de Asset Allocation se entiende como el paso inicial del proceso de inversión, el cual consiste en determinar cuanto del total de los fondos de la AFP se asigna a cada tipo de instrumento financiero<sup>25</sup>, para que luego los distintos managers gestionen los fondos asignados.

Se identifica como oportunidad de mejora que el proceso mencionado anteriormente posee una base objetiva “discrecional”, puesto que la decisión de la repartición del fondo se hace en base a un consenso al que llegan los gerentes de las distintas áreas involucradas con el gerente de asset allocation, donde toman en consideración múltiples indicadores financieros del tipo FTV<sup>26</sup> de una gran cantidad de activos financieros. Por lo tanto, en el proceso se pierde mucha información debido a que la interpretación y relación de las variables la realizan los managers.

Esto se relaciona directamente con los rendimientos de la AFP, puesto que si se coloca más dinero en los instrumentos que mejor performance tendrán, mayor es el resultado final obtenido. Además, la rentabilidad de la AFP tiene correlación positiva con la proporción del mercado que representa, más aún, se tiene correlación considerando rezagos, es decir, la rentabilidad final de un año se correlaciona, por ejemplo, a la cantidad de afiliados a fines del año siguiente (considerando 1 rezago). Lo anterior se muestra en la siguiente tabla:

<b>Rezago</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Correlación</b>	58%	51%	45%	35%

Tabla 4: Correlación entre participación de mercado y rendimiento<sup>27</sup>.

<sup>24</sup> France24. Quienes Protestan y Que Demandan. Disponible en: <https://www.france24.com/es/20191026-quienes-demandas-protestas-chile-radiografia> (consulta: 5 de enero de 2020).

<sup>25</sup> Renta fija local e internacional, renta variable local e internacional y FX.

<sup>26</sup> Fundamentales, técnicos y de valor.

<sup>27</sup> Elaboración propia.

Por lo tanto, se identifica como relevante desarrollar una metodología de asset allocation con base continua, es decir, que entregue resultados realizando un análisis conjunto de los datos donde estos hablen por sí solos, permitiendo capturar relaciones que no son directas del análisis “discrecional” y, de esta manera, tomar una mejor decisión. Lo anterior tiene directa relación con el retorno de la compañía, puesto que, como se menciona en el párrafo anterior, si se invierte más en aquellos instrumentos financieros que rendirán mejor en un periodo dado, mayor es la ganancia, más aún considerando que la cantidad de afiliados puede aumentar con una buena performance y, además, que estos en general cotizan al tope, puesto que el 6% del mercado de afiliados que posee la AFP representa el 19% de los activos transados al 31 de diciembre de 2018<sup>28</sup> <sup>29</sup>, implicando que los ingresos pueden ser aún mayores.

Para evidenciar el impacto de un aumento en la rentabilidad y en base a los resultados y rendimientos obtenido por la AFP en el año 2018 y publicados por la superintendencia de pensiones, un aumento de 10 puntos base<sup>30</sup> en la rentabilidad sobre el patrimonio inicial implica un incremento en los resultados de \$ 767.205.745 CLP<sup>31</sup>.

## 2.2. Hipótesis y Alternativas de Solución

La principal hipótesis del trabajo de título es la posibilidad de mejorar los resultados de la compañía utilizando una metodología estadística-matemática para la asignación de los fondos de la AFP, diferenciándose de la heurística actual utilizada por la compañía que implica un alto involucramiento de los managers.

A su vez, existe un abanico de herramientas que se pueden utilizar para generar la metodología solicitada. De estas se destacan métodos de aprendizaje de maquinas, puesto que son herramientas cuantitativas que se comportan bien con gran cantidad de datos, entregando buenos resultados si se manejan de buena forma. Por lo que se ajustan de buena manera al trabajo de título, donde se busca encontrar patrones, relaciones e información que no se evidencia a simple vista en la data.

---

<sup>28</sup> Superintendencia de pensiones. Afiliados AFP. Disponible en:

[https://www.spensiones.cl/apps/loadEstadisticas/genEstadAfiliadosCotizantes.php?id=inf\\_estadistica/aficot/mensual/2018/12/01C.html&p=M&menu=sci&menuN1=afil&menuN2=afp&orden=10&ext=.html](https://www.spensiones.cl/apps/loadEstadisticas/genEstadAfiliadosCotizantes.php?id=inf_estadistica/aficot/mensual/2018/12/01C.html&p=M&menu=sci&menuN1=afil&menuN2=afp&orden=10&ext=.html) (consulta: 1 de mayo de 2019).

<sup>29</sup> Superintendencia de pensiones. Activos AFP. Disponible en:

<https://www.spensiones.cl/apps/centroEstadisticas/paginaCuadrosCCEE.php?menu=sest&menuN1=sistpens&menuN2=fondospen> (consulta: 2 de mayo de 2019).

<sup>30</sup> Puntos base es una medida utilizada en finanzas para hablar sobre porcentaje, 100 puntos base = 1%. En adelante, puntos base es abreviado como PB.

<sup>31</sup> Superintendencia de pensiones. Centro de estadísticas. Disponible en:

<https://www.spensiones.cl/apps/centroEstadisticas/paginaCuadrosCCEE.php?menu=sci&menuN1=estfinafp&menuN2=NOID> (consulta: 9 de abril de 2019).

Existen cuatro tipos de aprendizaje de maquinas, los que según Gonzalo Alvaro (2018)<sup>32</sup> se definen como:

- Aprendizaje Supervisado: *“Estos métodos son los más sencillos de realizar. En ellos se parte de un **conocimiento a priori**. El objetivo es, mediante unos datos de entrenamiento, deducir una función que haga lo mejor posible el mapeo entre unas entradas y una salida. Los datos de entrenamiento constan de tuplas  $(X, Y)$ , siendo  $X$  las variables que predicen una determinada salida  $Y$ ”.*
- Aprendizaje no supervisado: *“Al contrario que en el aprendizaje supervisado, en este caso **no existe conocimiento a priori**. Aquí ya no se tienen tuplas  $(X, Y)$ , simplemente se tiene  $X$ . El objetivo del aprendizaje no supervisado es modelar la estructura o distribución de los datos para aprender más sobre ellos. Sirve tanto para entender, como para resumir un conjunto de datos”.*
- Aprendizaje semi supervisado: *“Lo que tenemos son tanto datos etiquetados como datos no etiquetados, es decir, además de tener tuplas  $(X, Y)$ , tenemos datos sólo de  $X$  de los que no sabemos su respuesta  $Y$ . El reto se encuentra en combinar datos etiquetados y no etiquetados para construir un modelo supervisado”.*
- Aprendizaje reforzado: *“El objetivo en el aprendizaje por refuerzo es aprender a mapear situaciones de acciones para maximizar una cierta función de recompensa. En estos problemas un agente aprende por prueba y error en un ambiente dinámico e incierto. En cada interacción el agente recibe como entrada un indicador de estado actual y selecciona una determinada acción que maximice una función de refuerzo o recompensa a largo plazo”.*

Por otro lado, el valor agregado que se genera al desarrollar la metodología impacta directamente a la rentabilidad de las inversiones de la AFP. Una forma de medir este impacto es mediante backtest, es decir, analizar los resultados que arroja el modelo en el pasado.

---

<sup>32</sup> Machine learning para todos. Tipos de aprendizaje automático. Disponible en: <http://machinelearningparatodos.com/tipos-de-aprendizaje-automatico/> (consulta: 30 de junio de 2019).

### **3. OBJETIVOS**

Se establecen los siguientes objetivos para el desarrollo del trabajo de memoria de título:

#### **3.1. Objetivo General**

Desarrollar una metodología de asset allocation que permita tomar una decisión objetiva de como distribuir la inversión entre los distintos tipos de instrumentos financieros para una AFP, todo esto con el fin de aumentar la rentabilidad de esta en base a modelos estadísticos/econométricos que utilizan información y datos provenientes de fuentes confiables como Bloomberg.

#### **3.2. Objetivos Específicos**

- Esquematizar el proceso de asset allocation para generar un entendimiento profundo del problema.
- Desarrollar modelos matemáticos-estadísticos que permitan determinar la asignación óptima para distribuir la inversión.
- Generar modelo de backtest para medir el desempeño de la metodología y, en caso de ser necesario, aplicar mejoras.



#### 4. ALCANCES

En primer lugar, se deben definir los tipos de activos con los que se trabaja. Se establecen los siguientes tipos de activos, que son en los que la compañía distribuye sus fondos en la actualidad y que, en el mediano plazo, no se modificarán:

- Renta variable mercado desarrollado: corresponden a instrumentos de renta variable como acciones de mercados desarrollados como Estados Unidos, Europa, etc.
- Renta variable mercado emergente: Mismos instrumentos que el punto anterior, pero en mercados emergentes como Latinoamérica, México, etc.
- Renta fija high yield: Corresponden a instrumentos de renta fija de mercados emergentes.
- Renta fija investment grade: Mismos instrumentos del punto anterior, pero de Estados Unidos.
- Renta fija local: Instrumentos de renta fija del mercado chileno.

Es importante aclarar que las monedas se utilizan para “protegerse” cuando se adquieren instrumentos en divisas volátiles o para cerrar posiciones, por lo que no se consideran al momento de generar el allocation.

Por otro lado, se desarrollan dos tipos de modelos que se basan en lo propuesto por Defend, Portelli, y Cesarini (2017) y que se mostrará como framework en el marco conceptual, los cuales son:

- Económico.
- Momentum.

A su vez, la metodología genera un portafolio óptimo que corresponde a la ponderación de los portafolios que genera cada estrategia por separado, es decir, se calculan dos posiciones óptimas para cada tipo de activo (una por cada modelo) y luego se ponderan.

El principal motivo para realizar el proyecto de esta manera viene dado por la ventana de tiempo en el que se debe desarrollar, se estima que calcular cada estrategia del framework mostrado en el marco conceptual excede el tiempo máximo, por lo que el óptimo es realizar el análisis de manera agregada con las técnicas seleccionadas y, así, perder la menor cantidad de representatividad posible y alcanzar a realizar el proyecto en los tiempos estipulados.

## 5. RESULTADO DE LA METODOLOGÍA

Se tiene como resultado de la metodología la construcción de un portafolio con la distribución de los fondos para cada tipo de instrumento financiero con periodicidad mensual. Tal como se expone en el siguiente ejemplo:

<b>1-jul</b>	<b>Económico</b>	<b>Momentum</b>	
<b>Ponderador (%)</b>	<b>50%</b>	<b>50%</b>	<b>Allocation</b>
<b>R.V. Mercado Desarrollado</b>	0.15	0.4	<b>0.275</b>
<b>R.V. Mercado Emergente</b>	-0.15	0.1	<b>-0.025</b>
<b>R.F. High Yield</b>	0.2	-0.25	<b>-0.025</b>
<b>R.F. Investment Grade</b>	-0.1	-0.15	<b>-0.125</b>
<b>R.F.L</b>	-0.2	-0.1	<b>-0.15</b>
<b>Cash</b>	0.1	0	<b>0.05</b>

Tabla 5: Allocation final.

Como es posible evidenciar, el portafolio de cada estrategia es ponderado para determinar el allocation final. El mánager debe mover los fondos de clase de activo en clase de activo con el fin de aumentar las ganancias de la compañía.

El rebalanceo es mensual debido a que los datos que se utilizan para desarrollar unas de las estrategias tiene esta periodicidad. Por lo que no tiene sentido realizarlo con una periodicidad menor, ya que no capturaría los nuevos efectos en la economía y el mercado.

Por último, es importante destacar que el allocation es un portafolio de suma cero, el cual representa la variación con respecto al allocation actual. Esta variación es la que, en teoría, se debe realizar cada uno de los cinco fondos existentes, sin embargo, estos tienen restricciones por lo que es posible que en ocasiones no se pueda realizar tal modificación, principalmente en los fondos más restrictos como el E.

## 6. MARCO TEÓRICO

### 6.1. Framework

El marco teórico es fundamentalmente guiado por publicaciones de distintos managers de inversión como Amundi, JP Morgan, Robeco, etc. Dentro de estos se destaca lo realizado por Defend, Portelli and Cesarini (2017) que se toma como framework para el trabajo de título, el cuál se detalla a continuación:

	Strategy (Based on Analytical Tool)	ACA Pillar	Style	Reference Universe	Horizon	Type	Techniques
1	Advanced Phazer	Ec backdrop	Growth	Cross Asset	Strategic	Directional Macro	K-means algorithm based on Internal macro forecast
2	Macro Target Prices	Ec. backdrop	Growth/Value	Cross Asset	Strategic	Directional Macro	Cointegration based on Internal macro forecast
3	Economic Surprise	Ec backdrop	Momentum	Global Equities	Tactical	Relative	Filtering
4,a	Valuation (KNN)	Valuation	Value	Global Equities	Strategic	Directional/Relative	K-Nearest Neighbors algorithm & Internal macro forecast
4,b	Valuation (SVM)	Valuation	Value	Global Equities	Strategic	Directional/Relative	Vector machine algorithm & Internal macro forecast
4,c	Valuation (dt)	Valuation	Value	Global Equities	Strategic	Directional/Relative	Random Forest algorithm & Internal macro forecast
5	Technical Analysis	Technicals	Momentum	Cross Asset	Tactical	Directional Sentiment	Filtering
6	Flows	Technicals	Momentum	Cross Asset	Tactical	Directional Sentiment	Filtering
7	RSI	Technicals	Momentum	Cross Asset	Tactical	Directional/Relative	Filtering
8	Risk On Risk Off (RORO)	Risk Sentiment	Volatility	Cross Asset	Tactical	Directional Sentiment	PCA, K means algorithm

Figura 3: Framework Asset Class Assessment.

El framework entrega ocho estrategias que miden distintas variables que afectan a los activos financieros, como crecimiento económico, momentum, value y sentimiento de riesgo. A su vez, para cada estrategia se proponen técnicas para capturar el efecto de las variables. Si bien no se detalla como realizar cada una, sirve como guía para plantear soluciones.

Momentum se entiende como una medida que captura la tasa de aceleración del precio de un activo<sup>33</sup>. Ya sea a través del precio en sí (por ejemplo, para acciones) o a través de tasas (por ejemplo, para bonos). En simples palabras, captura la tendencia de un activo financiero.

<sup>33</sup> Investopedia. Momentum. Disponible en: <https://www.investopedia.com/terms/m/momentum.asp> (consulta: 30 de junio de 2019).

Value se entiende como una medida que captura la relación entre el valor actual del activo con el valor intrínseco de este. En simples palabras, captura el hecho de si el activo está caro o barato.

Crecimiento económico hace relación al mercado, y como este afecta en el proceso de inversión, implicando que los distintos activos financieros se comporten de distintas maneras bajo ciertas condiciones.

Sentimiento de riesgo se entiende como una medida del sentimiento del mercado, en el sentido de estimar la exposición óptima bajo ciertos estándares. En simples palabras, determina si es prudente arriesgarse, y que tanto riesgo se debe tomar en el proceso de inversión.

## **6.2. Optimización de Portafolios**

Según lo expuesto, cada estrategia debe generar un portafolio distinto en base a las variables que considera. Por lo que se generan dos interrogantes, en primer lugar, como construir el portafolio óptimo para cada estrategia y, en segundo lugar, como ponderar cada portafolio para construir la recomendación final.

Para resolver la primera interrogante se tiene lo expuesto por Bourgeron, Lezmi y Roncalli (2018), donde analizan la teoría de portafolio moderno propuesto por Markowitz en 1952 y que actualmente es usada por managers de inversión<sup>34</sup>, a su vez, estudia variaciones de este donde se agregan restricciones que lo hacen más robusto y menos sensible ante variaciones en las condiciones iniciales. Por otro lado, se tiene lo expuesto por Husnain, Hassan y Lamarque (2016), donde evidencian que métodos complejos para mejorar el resultado de la teoría de Markowitz no tienen un mejor performance que los más simples. Por último, se tiene lo expuesto por Lim, Roberts y Zohren (2019), donde proponen un método alternativo para mejorar los resultados de Markowitz mediante un ajuste a la matriz de varianza-covarianza.

Para resolver la segunda interrogante se tiene lo expuesto por Defend, Portelli y Cesarini (2017), donde plantean utilizar el sharpe ratio<sup>35</sup> para determinar que tanto pondera cada estrategia en la recomendación final. También se puede considerar lo expuesto por Mallard (2018), donde propone ajustar el sharpe ratio al riesgo de eventos extremos, característica que no se captura por sí sola. Sin embargo, este desarrollo queda propuesto para otro trabajo de investigación, puesto que sale del foco principal de este proyecto de título.

---

<sup>34</sup> Investopedia. Modern Portfolio Theory. Disponible en: <https://www.investopedia.com/managing-wealth/modern-portfolio-theory-why-its-still-hip/> (consulta: 28 de agosto de 2019).

<sup>35</sup> Su valor indica el premio por unidad de riesgo. Se define como la diferencia entre el retorno esperado del portafolio y el activo libre de riesgo dividido por la volatilidad del portafolio.

### 6.3. Estrategia Económica

Para esta estrategia Defend, Portelli y Cesarini (2017) proponen realizar clustering de variables macroeconómicas utilizando la técnica de k-means (en adelante KM), esto con el objetivo de determinar valores de referencia para cada fase identificada de un ciclo económico, para luego calcular las probabilidades de estar en cada fase del ciclo económico en el periodo actual. La probabilidad se calcula en base a la inversa de la distancia euclidiana entre los valores de las variables macroeconómicas actuales y las de referencia de cada fase de un ciclo, por lo tanto, entre menor es la distancia, mayor es la probabilidad de estar en una fase del ciclo. A su vez, se propone la técnica de modelos gaussianos mixtos (en adelante GMM), que es alternativa a k-means y posee el mismo objetivo mencionado, esta considera la distribución de los datos para determinar probabilidades de pertenencia a los clusters mediante un algoritmo de maximización de expectativas<sup>36</sup>. Para finalizar, se debe determinar el portafolio óptimo en base a las probabilidades de estar en cada fase (tanto para GMM y para KM).

#### 6.3.1. Clustering: KM y GMM

El clustering<sup>37</sup> es una técnica no supervisada de aprendizaje de máquinas, por lo tanto, busca generar entendimiento sobre la distribución o estructura de los datos. Su objetivo es dividir los puntos de datos en varios grupos, de manera de que dentro de cada grupo los datos sean similares entre sí y fuera del grupo sean distintos.

K-means es una técnica de clustering que funciona de la siguiente manera:

- Se le ingresa como parámetro la cantidad de clusters que se quieran generar.
- K-means genera centroides<sup>38</sup> aleatorios.
- Cada n-tupla de datos que representa una observación es asignada al cluster más cercano.
- Los centroides se van iterando de manera que se vaya minimizando la suma de la distancia euclidiana de cada n-tupla de datos con el centroide más cercano y se maximice la distancia entre los grupos.

---

<sup>36</sup> Explicado de manera simple, este algoritmo busca maximizar la probabilidad de encontrar la data clusterizada en cada cluster.

<sup>37</sup> Geeks For Geeks. Clustering in Machine Learning. Disponible en :<https://www.geeksforgeeks.org/clustering-in-machine-learning/> (consulta:29 agosto 2019).

<sup>38</sup> Vector en un espacio n-dimencional que representa el centro de un grupo. N corresponde a la cantidad de variables.

GMM es una técnica de clustering que funciona de la siguiente manera:

- Se ingresa como parámetro la cantidad de clusters que se quieren generar.
- GMM genera centroides aleatorios.
- Cada punto es asignado a cada uno de los clusters con diferentes probabilidades.
- Los centroides se van iterando de manera que se maximice la verosimilitud del problema estudiado, es decir, se maximice la probabilidad de encontrar esa data en cada cluster.

Para GMM se asume que los datos distribuyen normal, lo que no es un impedimento dadas las características del problema, puesto que según el teorema central del límite, las medias de muestras grandes ( $n > 50$ ) y aleatorias son aproximadamente normales<sup>39</sup>, características presentes en la data a estudiar. A su vez, es interesante el hecho de que al distribuir normal y existir matrices de varianza covarianza asociadas a los clusters, la asignación toma en consideración las correlaciones entre las variables.

Como es posible deducir, una de las principales diferencias entre las técnicas es el hecho de que KM determina una pertenencia absoluta de la entrada de data al cluster y GMM determina una pertenencia relativa en base a una probabilidad, esta característica se entiende como asignación fuerte y asignación suave respectivamente<sup>40</sup>. Por lo tanto, lo que se realiza con KM es suavizar la asignación, creando un criterio para determinar una probabilidad.

Se adjunta en anexo 1 el desarrollo teórico de cada algoritmo de clustering.

---

<sup>39</sup> Support Minitab. Sobre el Teorema Central del Límite. Disponible en: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/data-concepts/about-the-central-limit-theorem/> (consulta: 31 diciembre 2019).

<sup>40</sup> SandiPanWeb. Hard and Soft Clustering with K-means and GMM. Disponible en: <https://sandipanweb.wordpress.com/2017/03/19/hard-soft-clustering-with-k-means-weighted-k-means-and-gmm-em/> (consulta: 31 diciembre 2019).

## 6.4. Estrategia Momentum

Para esta estrategia Defend, Portelli y Cesarini (2017) proponen tratar estadísticamente los datos a utilizar, en este sentido, se considera lo expuesto por Lim, Roberts y Zohren (2019) donde utilizan una estrategia propuesta por Baz et al. (2015) para distintas escalas de tiempo. El objetivo de esto es determinar un portafolio óptimo en base a la tendencia que tienen los activos mediante la técnica de media móvil convergencia/divergencia (en adelante MACD).

### 6.4.1. MACD

Según Investopedia<sup>41</sup>, MACD se entiende como un indicador de tendencia que muestra la relación entre dos promedios móviles (de distintos periodos) del precio de un mismo activo. A su vez, este indicador tiene como característica la capacidad de mostrar la duración y los cambios en la dirección de una tendencia del precio de un activo, entre otras cosas<sup>42</sup>.

La relación entre los promedios móviles para un activo “i” en tiempo “t” es la siguiente:

$$MACD(i, t, S, L) = m(i, S) - m(i, L)$$

Donde  $m(i, S)$  corresponde a un EWMA<sup>43</sup> de los precios del activo “i” con escala de tiempo S, que se traduce en una vida media<sup>44</sup> (HL) de  $\frac{\log(0.5)}{\log(1-\frac{1}{S})}$ . Es importante destacar el hecho de que  $S < L$  para que MACD sea capaz de capturar la tendencia.

En este sentido, se tiene que la combinación de dos promedios móviles muestra el cambio en el impulso<sup>45</sup> de una tendencia<sup>46</sup>, por lo tanto, MACD es capaz de capturar esta característica.

---

<sup>41</sup> Investopedia. MACD. Disponible en: <https://www.investopedia.com/terms/m/macd.asp> (consulta: 2 de enero de 2020).

<sup>42</sup> Tradeciety. How to Use the MACD. Disponible en: <https://www.tradeciety.com/tips-how-to-use-the-macd/> (consulta: 2 de enero de 2020).

<sup>43</sup> Exponential Weighted Moving Average. Corresponde a una técnica de series de tiempo que, en simple palabras, es un promedio móvil ponderando que otorga distinta relevancia a la historia de una data.

<sup>44</sup> StackExchange. Half Life of EWMA. Disponible en: <https://quant.stackexchange.com/questions/46194/half-life-of-exponential-weighted-moving-average> (consulta: 2 de enero de 2020).

<sup>45</sup> Entiendase impulso como la medida de que tanto va a seguir durando la tendencia, es decir, a mayor impulso más se espera que esa tendencia continúe.

<sup>46</sup> Tradeciety. How to Use Moving Averages. Disponible en: <https://www.tradeciety.com/how-to-use-moving-averages/> (consulta: 2 de enero de 2020).

Por otro lado, la condición “convergente/divergente” hace relación a una característica de este indicador, por ejemplo, se habla de divergencia cuando MACD presenta dos mínimos ascendentes en su valor y, a su vez, el precio del activo presenta dos mínimos descendientes, puesto que la pendiente en MACD es negativa y en el precio es positiva. Esta divergencia se denomina alcista, puesto que prevé que el precio va a subir, como se aprecia en la siguiente figura:



Figura 4: Criterio de Divergencia Alcista<sup>47</sup>.

<sup>47</sup> Investopedia. MACD. Disponible en: <https://www.investopedia.com/terms/m/macd.asp>. (Consulta: 2 de enero de 2020).



## 7. METODOLOGÍA

La metodología está enfocada en la ciencia de los datos, puesto que contempla varias características de esta, entre las que se destaca el conocimiento o dominio del problema, análisis y tratamiento de datos, desarrollo de modelos y evaluación de estos. Todo lo anterior puede realizarse utilizando software como Python y datos de fuentes como Bloomberg. La metodología se puede resumir en la siguiente figura:

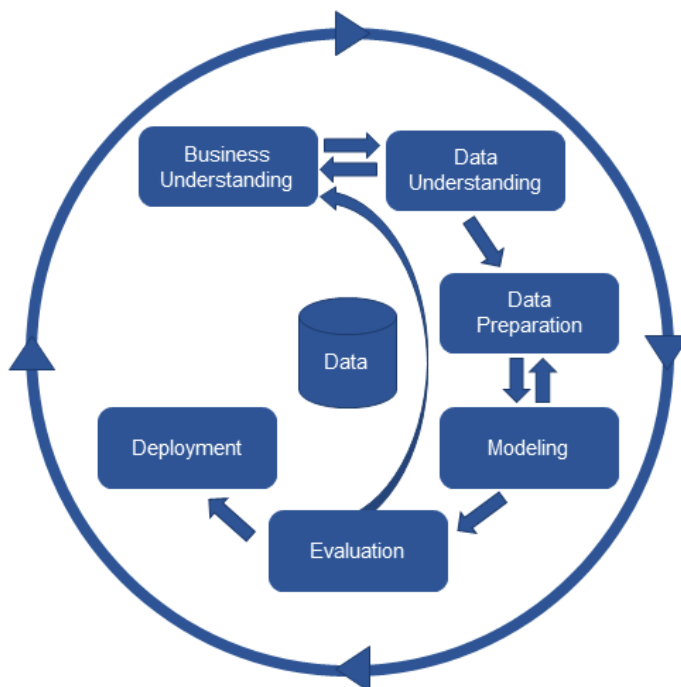


Figura 5: Metodología de la Ciencia de los Datos.

Este proceso se relaciona de manera directa con lo expuesto en el marco conceptual, puesto que toma en consideración cada una de las partes que se involucran con la bibliografía propuesta.

El dominio del problema es el punto de partida para determinar que propuesta de cada una de las estudiadas entrega mejores resultados para la oportunidad de mejora encontrada y, de esta manera, diseñar otra que sea más robusta y se ajuste a las condiciones del trabajo de título.

Por otro lado, el análisis y tratamiento de los datos es un punto sumamente relevante, puesto que es la base del desarrollo del trabajo de título. Se destaca que las fuentes a las que tiene acceso la compañía son confiables, tal como se expone en Investopedia<sup>48</sup>: “*Bloomberg es el mayor proveedor de noticias e información financiera a nivel global, incluyendo datos en tiempo real e históricos*”.

A su vez, el desarrollo de modelos es el siguiente paso para obtener el resultado final que, tal y como se expone en el marco conceptual, implica utilizar algoritmos del aprendizaje de maquina supervisado y no supervisado para los distintos tipos de estrategias, como también modelos de optimización como mean-variance (Markowitz). La principal razón de utilizar estos algoritmos viene de que son propuestos y utilizados por grandes managers de inversión, comportándose de buena manera para el trabajo de título.

También se destaca que para el desarrollo de modelos y el tratamiento de datos se utiliza Python, lenguaje de programación que ha ido adquiriendo notoria popularidad en el último tiempo, siendo uno de los más utilizados y que, además, cuenta con más de 145.000 paquetes de software personalizados para distintos usos, donde se destacan Numpy, Pandas y Seaborn para el tratamiento de datos y Sklearn para algoritmos del aprendizaje de maquinas<sup>49</sup>.

Con todo lo anterior en consideración, se define la siguiente metodología para el proyecto:

- Estudiar y entender cada estrategia a desarrollar.
- Seleccionar data necesaria.
- Depurar data.
- Desarrollar modelo matemático-estadístico propio de la estrategia.
- Testear y mejorar el modelo.
- Implementar estrategia.

---

<sup>48</sup> Investopedia. Bloomberg. Disponible en: <https://www.investopedia.com/terms/b/bloomberg.asp> (consulta: 30 de junio de 2019).

<sup>49</sup> Azul web. Python lenguaje más popular del mundo. Disponible en: <https://www.azulweb.net/python-se-esta-convirtiendo-en-el-lenguaje-de-programacion-mas-popular-del-mundo/> (consulta: 30 de junio de 2019).

## 8. DESARROLLO DEL PROYECTO

### 8.1. Índices

En primer lugar, es importante identificar los índices que representan las clases de activos en los que se va a realizar allocation. Estos son utilizados por los modelos desarrollados y, a su vez, definidos y entregados por el mánager de inversión:

- LVACLI Index: Cash local.
- LD20TRUU: Cash USD.
- .DMEQLCEW: Renta variable mercados desarrollado (RVD).
- .EMEQCULC: Renta variables mercados emergentes(RVE).
- LUATTRUU: Renta fija investment grade US (RFIG).
- .CHFICUAA. Renta fija local (RFL).
- .HBFIWOCO: Renta fija high yield<sup>50</sup> (RFHY).

Es importante recalcar que la monedas no se incluyen, en primera instancia, en el proceso de allocation, puesto que como se menciona en el alcance, se utilizan para protegerse de divisas volátiles o cerrar posiciones.

---

<sup>50</sup> Andbank. ¿Qué significa High Yield?. Disponible en: <https://www.andbank.es/observatoriodelinversor/que-significa-high-yield/> (consulta: 17 de octubre de 2019).

## 8.2. Modelo Económico

Para el desarrollo de este modelo se utiliza lo publicado por Defend, Portelli and Cesarini (2017) en “Investment Phazer”.

### 8.2.1. Data

La data utilizada es la propuesta en la publicación, donde se considera como punto de partida variables macroeconómicas de Estados Unidos, agrupándose de la siguiente manera: crecimiento, inflación, política monetaria y deuda. Por otro lado, es importante destacar que la data comienza desde el 1 de enero de 2000 hasta el 31 de mayo de 2019 y presenta periodicidad mensual, el tamaño de la muestra se determina en base al conocimiento experto, donde se concluye que es un periodo en el cual se puede capturar un ciclo completo. Es importante recalcar que las variables son seleccionadas junto a un experto según las características que poseen, puesto que todas las variables tienen características leading -a excepción de la inflación que presenta rezagos-, es decir, son precursoras ante cambios en el mercado y no poseen rezagos.

Por otro lado, la operación se realiza el tercer lunes de cada mes, puesto que en esa fecha se tiene con seguridad todas las variables seleccionadas. Por ejemplo, el 17 de junio se realizaría el allocation con la data correspondiente al 31 de mayo.

#### Crecimiento

Para el crecimiento se utilizan las siguientes variables:

- PMI manufacturero<sup>51</sup>: Es una medida macroeconómica que intenta capturar la situación económica de un país a través de una encuesta que se les realiza a las empresas más representativas del mercado.
- EPS: Representan las ganancias por acción de las empresas más representativas del mercado de US.
- SPS: Representa los ingresos por acción de las empresas más representativas del mercado.

---

<sup>51</sup> Wikipedia. Purchasing Management Index. Disponible en: [https://es.wikipedia.org/wiki/Purchasing\\_Managers\\_Index](https://es.wikipedia.org/wiki/Purchasing_Managers_Index) (consultado: 29 de agosto de 2019).

Además, se determinan las siguientes variables:

- Tasa de crecimiento de ganancias:  $GE_t = \frac{E_t}{E_{t-1}} - 1$
- Tasa de crecimiento de ingresos:  $GS_t = \frac{S_t}{S_{t-1}} - 1$
- Margen:  $M_t = \frac{E_t}{S_t}$

Es importante destacar que no se utiliza el GDP (PIB) debido a que tiene una periodicidad de 3 meses, implicando que se pierda mucha información para el desarrollo del modelo.

### Inflación

Para la inflación se utiliza como variable el índice de precios del consumidor, que mide en promedio como cambian los precios pagados.

### TPM

Para la tasa de política monetaria se utiliza como variable la tasa de política monetaria de US.

### Deuda

Con respecto a la deuda, se toman como variables las relaciones: deuda total/total assets, deuda neta/EBITDA, deuda total/total equity y deuda neta per share de las empresas más relevantes del mercado de US.

En resumen, los datos a utilizar en este modelo son los siguientes:

<b>Tipo Data</b>	<b>Data</b>
Índices	.DMEQLCEW Index
	.EMEQCULC Index
	.CHFICUAA Index
	LUATTRUU Index
	.HBFIWOCO Index
Inflación	CPI YOY Index
Crecimiento	PMI Manu
	Earnings
	Sales
	Earnings Growth
	Sales growth
	Margin
TPM	TPM Hoy
Deuda	Deuda_total/Total Assets
	Deuda_Neta/EBITDA
	Deuda_total/Total Equity
	Deuda neta per share

Tabla 6: Data total.

### 8.2.2. Tratamiento de Datos

Se realiza un análisis exploratorio de datos en busca de valores faltantes, correlaciones y outliers que puedan afectar al resultado final del proyecto, obteniéndose los resultados mostrados a continuación. Este es un paso clave para el desarrollo del modelo, puesto que es lógico pensar que independiente de que tan bien esté construido o formulado, si se ingresan valores erróneos, erróneos son los resultados.

#### Datos Faltantes

Se busca la existencia de datos faltantes para verificar la calidad de la data, en el sentido de que esta tenga consistencia a través del tiempo y verificar que se actualice periódicamente. A su vez, la existencia de datos faltantes puede generar problemas en el desarrollo del modelo, implicando que este no se pueda ejecutar de buena manera.

<b>Index</b>	<b>Datos Faltantes</b>
CPI YOY Index	0
PMI Manu	0
Earnings Pasados	0
Sales Pasados	0
Past Earnings growth rate	0
Past Sales growth rate	0
Margin Pasado	0
TPM Fed Hoy	0
Deuda Total/Total Assets	0
Deuda Neta/EBITDA	0
Deuda Total/Total Equity	0
Deuda Neta per Share	0
.DMEQLCEW Index	0
.EMEQCULC Index	0
.CHFICUAA Index	12
LD20TRUU Index	0
LVACLI Index	12
LUATTRUU Index	0
.HBFIWOCO Index	23

Tabla 7: Datos Faltantes.

Se destaca que la dato macro no presenta datos faltantes. Por otro lado, los datos faltantes correspondientes al índice representativo de la renta fija local, cash local y crédito extranjero global corresponden a los datos iniciales de la serie de tiempo, puesto que no existe información en esa fecha, sin embargo, no se eliminan ni se reemplazan de la data debido a que no afecta de gran manera al resultado final.

### Outliers

Se realizan boxplots para detectar la presencia de outliers en la data a utilizar. Estos pueden generar distorsión en los resultados, puesto que se identifican como valores (típicamente erróneos) muy distintos al resto que influyen en la media de los datos, implicando que el resultado final del modelo no esté bien calibrado.

Para la data macroeconómica se tiene:

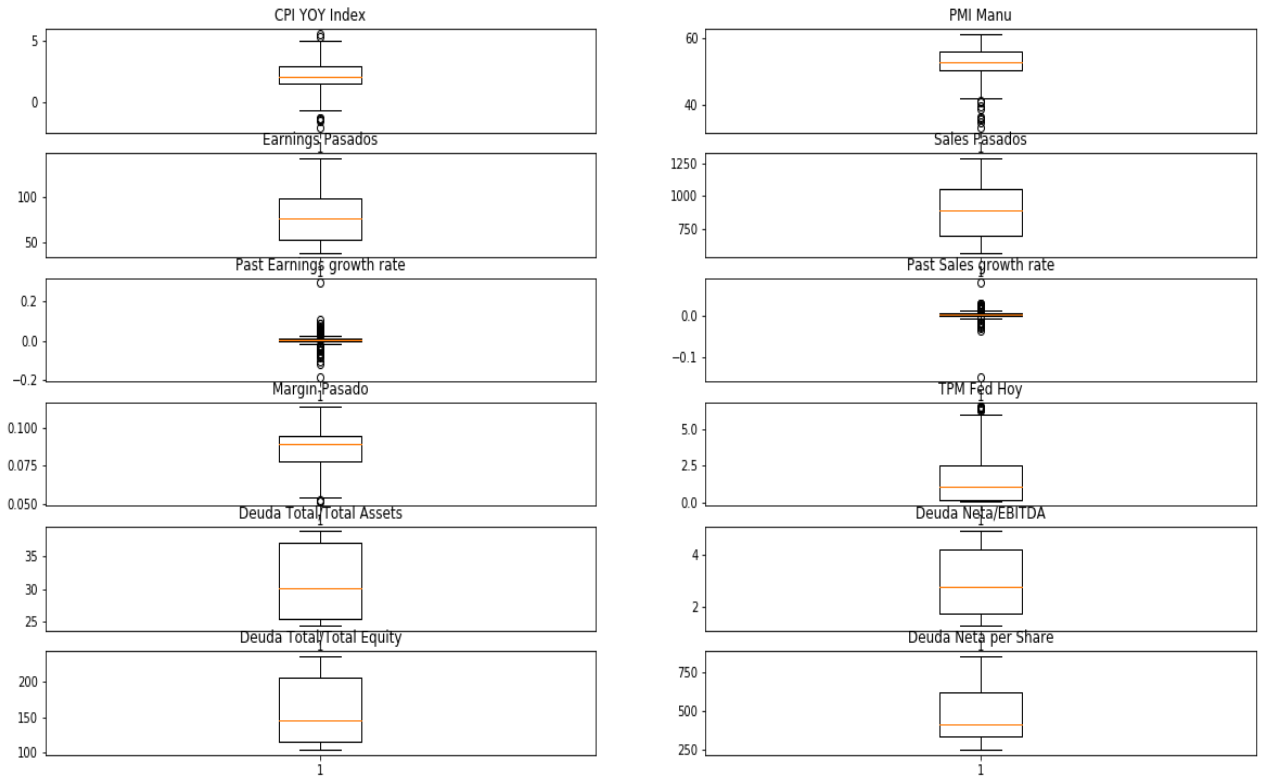


Figura 6: BoxPlot Data Macroeconómica.

Para los índices se tiene:

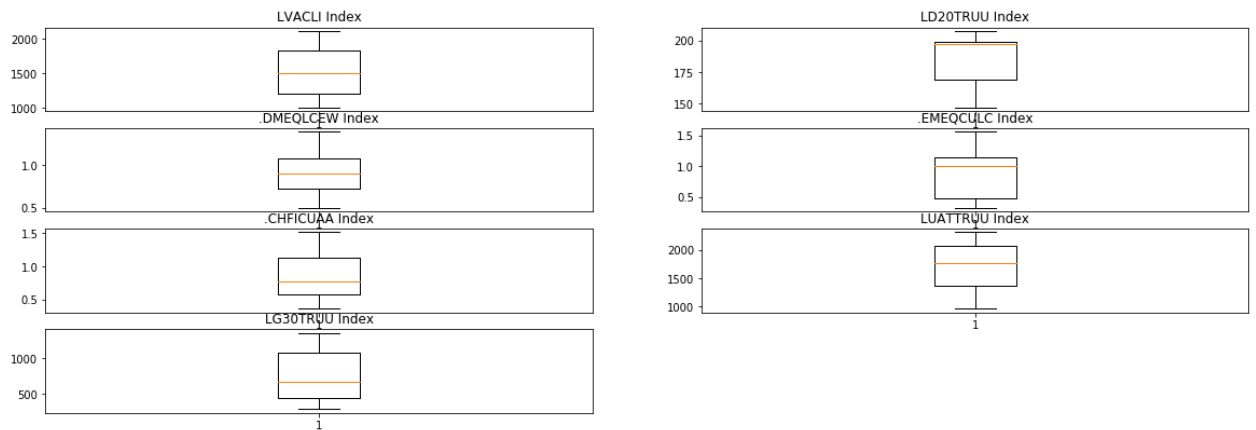


Figura 7: BoxPlot Índices Representativos.



El análisis se realiza junto al jefe de asset allocation de la AFP, concluyéndose que:

- Es esperable grandes variaciones en las tasas de crecimiento debido a su composición.
- Los eventos que parecen outliers en la variable de deuda, PMI e inflación pueden ser provocados por shocks inesperados como crisis.
- Los índices no presentan outliers.

Por lo tanto, no se elimina ninguna variable.

### Correlación

Se realiza un análisis de correlación de las variables macroeconómicas, puesto que se utilizan para generar los clusters. El objetivo de esto es no utilizar variables que estén altamente correlacionados con otras para no complejizar innecesariamente el modelo y, a su vez, no caer en la maldición de la dimensionalidad<sup>52</sup>. Se obtiene el siguiente gráfico (se adjunta en anexo 2 la matriz de correlación):

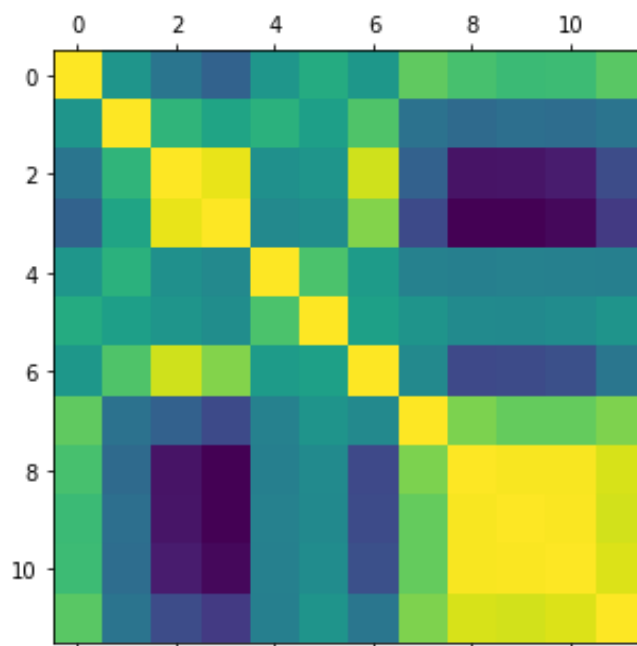


Figura 8: Correlaciones entre Data Macroeconómica.

<sup>52</sup> Ingeniero Beta. Maldición de la Dimensionalidad. Disponible en: <http://ingenierobeta.com/maldicion-de-la-dimensionalidad/> (consulta: 15 de octubre de 2019).

Para entender la figura es importante aclarar que entre más claro es el color, mayor es la correlación. Por otro lado, si dos variables se encuentran muy correlacionadas implica que agregar ambas al modelo no genera valor puesto que se mueven a la par.

Del análisis se concluye que:

- Las variables de deuda tienen alta correlación entre sí.
- EPS y SPS se encuentran altamente correlacionadas.

Por lo tanto, se decide utilizar como variable de deuda solo la “debt per share”, puesto que, en general, presenta menos correlación con las otras variables. Además, se decide utilizar EPS sobre SPS debido a que según el criterio del mánager las ganancias son más representativas para explicar un ciclo económico.

Entonces, para desarrollar el modelo la data final que se utiliza es la siguiente, la cual tiene 233 entradas:

<b>Tipo Data</b>	<b>Data</b>
Índices	.DMEQLCEW Index
	.EMEQCULC Index
	.CHFICUAA Index
	LUATTRUU Index
	.HBFIWOCO Index
Inflación	CPI YOY Index
Crecimiento	PMI Manu
	Earnings
	Earnings Growth
	Sales Growth
	Margin
TPM	TPM Hoy
Deuda	Deuda Neta Per Share

Tabla 8: Data Final.

### 8.2.3. Clustering de Fase y Probabilidad de Pertenencia

Para ambos clustering se utiliza la data macro formada por las variables de inflación, crecimiento, TPM y deuda ya depurados.

#### Clustering: K-Means

Como se menciona en el marco teórico, el algoritmo de k-means necesita como parámetro la cantidad de clusters que se quieren generar, por lo tanto, hay que determinar la cantidad óptima para el modelo mediante el método del codo.

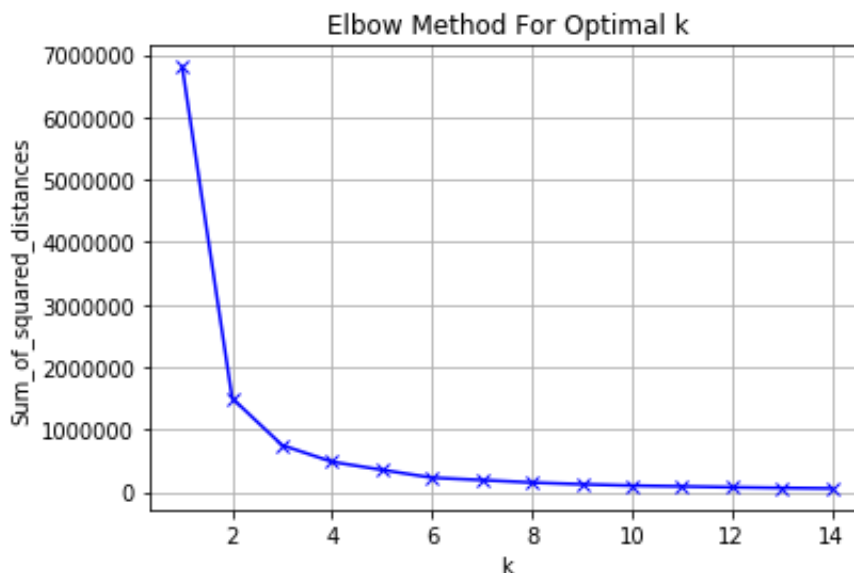


Figura 9: Método del Codo.

El gráfico muestra la suma de las distancias al cuadrado de todos los puntos con su respectivo cluster. Como es de esperar, a medida que la cantidad de clusters aumenta menor es la suma de las distancias. Por otro lado, el óptimo se encuentra en donde el cambio de pendiente es mayor (codo), puesto que es el punto donde más valor agrega el generar un nuevo cluster, además, hay que tener en consideración que agregar más clusters complejiza el problema de optimización propio del algoritmo.

De la figura 9 se desprende que el óptimo se encuentra entre 3 y 6 clusters, lo cual se considera como un buen resultado ya que, en general, se define que un ciclo económico tiene 4 fases<sup>53</sup>. Entonces, se escoge como óptimo  $k = 4$ . Cada centroide se entiende como los valores de referencia de las variables para cada una de las 4 fases (en anexo 3 se adjuntan los centroides estimados).

<sup>53</sup> Medium. Conoce las 4 Fases del Ciclo Económico. Disponible en: <https://medium.com/@facilitadorimpulsa/conoce-las-4-fases-del-ciclo-econ%C3%B3mico-53a904d26b47> (consulta: 29 de agosto de 2019).

## Asignación de Probabilidades: K-Means

La estrategia debe asignar una probabilidad de pertenencia de cada entrada de datos a cada una de las fases (clusters). Se define la probabilidad de que una entrada de datos "i" pertenezca a la fase "j" según lo que se subentiende de Defend, Portelli and Cesarini (2017) de la siguiente manera:

$$Prob_{i,j} = \frac{\frac{1}{dis_{i,j}}}{\sum_k \frac{1}{dis_{i,k}}} ; k: clusters$$

Donde  $dis_{i,j}$  corresponde a la distancia euclidiana entre el centroide j y la entrada i. Se obtiene la siguiente evolución de las probabilidades (cada color es una fase):

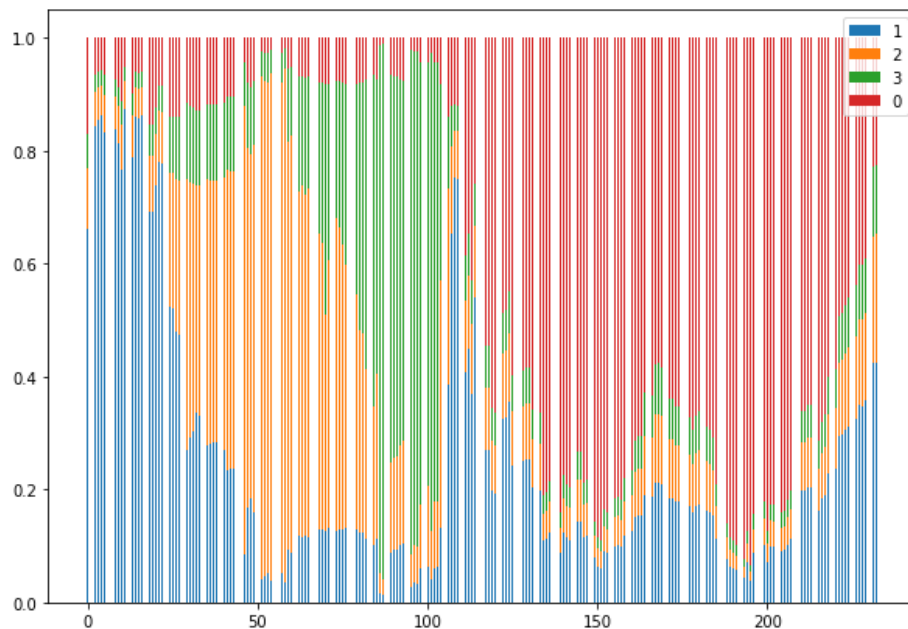


Figura 10: Evolución de las Probabilidades KM.

De la figura 10 es posible apreciar que la estrategia es capaz de identificar momentos en que es altamente probable que se esté en una fase en específico, tomando fase 1, fase 2, fase 3 y, por último, fase 0, que abarca una gran cantidad de tiempo. Por lo tanto, es posible inferir que el algoritmo es capaz de capturar la diferencia en los datos a partir de los parámetros entregados.

## Clustering: GMM

Al igual que con k-means, es necesario entregar como parámetro la cantidad de clusters que se desean generar, para el caso de esta técnica se determina el número óptimo de clusters utilizando como medida el BIC:

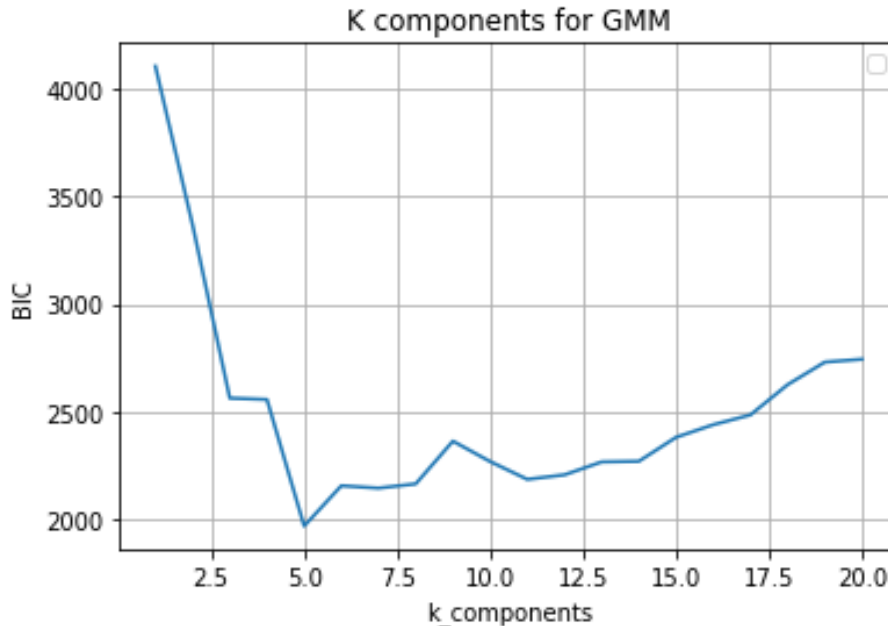


Figura 11: Criterio BIC.

Esta medida entrega un valor de que tan bueno es el modelo GMM para predecir la data que actualmente se tiene dado los clusters que se eligen. De la siguiente manera:

$$BIC = -2 * \ln(\hat{L}) + k * \ln(n)$$

Con  $\hat{L}$ : Maximo valor de la función de verosimilitud del modelo<sup>54</sup>.

Como es posible apreciar, el primer termino de la formula captura que tan bueno es el modelo y, el segundo termino penaliza por la cantidad de clusters que se eligen. Por lo tanto, mejor es el modelo a medida que disminuye el valor del BIC. De la figura 11 es posible apreciar que donde ocurren los mayores cambios de pendiente es entre 3 y 5 clusters, en este sentido, se escoge como óptimo  $k = 5$ , la razón de esto es que, en primer lugar, es el punto donde se genera el mayor cambio de pendiente y, en segundo lugar, sigue la línea de modelos exóticos expuesta por Defend, Portelli and Cesarini (2017) con una quinta fase identificada.

<sup>54</sup> Wikipedia. Criterio BIC. Disponible en: [https://es.wikipedia.org/wiki/Criterio\\_de\\_informaci%C3%B3n\\_bayesiano](https://es.wikipedia.org/wiki/Criterio_de_informaci%C3%B3n_bayesiano) (consulta: 8 de enero de 2020).

## Asignación de probabilidades: GMM

Tal y como en KM, la estrategia debe asignar una probabilidad de pertenencia de cada entrada de datos a cada una de las fases definidas. Dadas las características de este modelo, no se debe idear ninguna fórmula especial ya que las probabilidades son calculadas para la asignación a los clusters en la formulación del algoritmo. Dicho esto, la probabilidad de que la entrada de datos "i" ( $x_i$ ) pertenezca a la fase "j" viene dada por:

$$Prob_{i,j} = \frac{(\phi_j * N(x_i; \mu_j, \Sigma_j))}{\sum_k \phi_k * N(x_i; \mu_k, \Sigma_k)} ; k: clusters$$

Donde:

$$\phi_j = \frac{1}{N} * \sum_i Prob_{i,j} ; \mu_j = \frac{\sum_i Prob_{i,j} * x_i}{\sum_i Prob_{i,j}} ; \Sigma_j = \frac{\sum_i Prob_{i,j} * (x_i - \mu_j) * (x_i - \mu_j)^T}{\sum_i Prob_{i,j}}$$

Con:

$N()$ : Distribución normal (Gauss).

$\phi_j$ : Peso asociado al cluster j, es decir, representa la "fuerza" del gaussiano j.

$\mu_j$ : Media del cluster j.

$\Sigma_j$ : Matriz de varianza-covarianza del cluster j.

Luego, para determinar la probabilidad óptima de pertenencia se utiliza el algoritmo de maximización de expectativa que se encuentra adjunto en anexo 1. Obteniéndose la siguiente evolución de probabilidades:

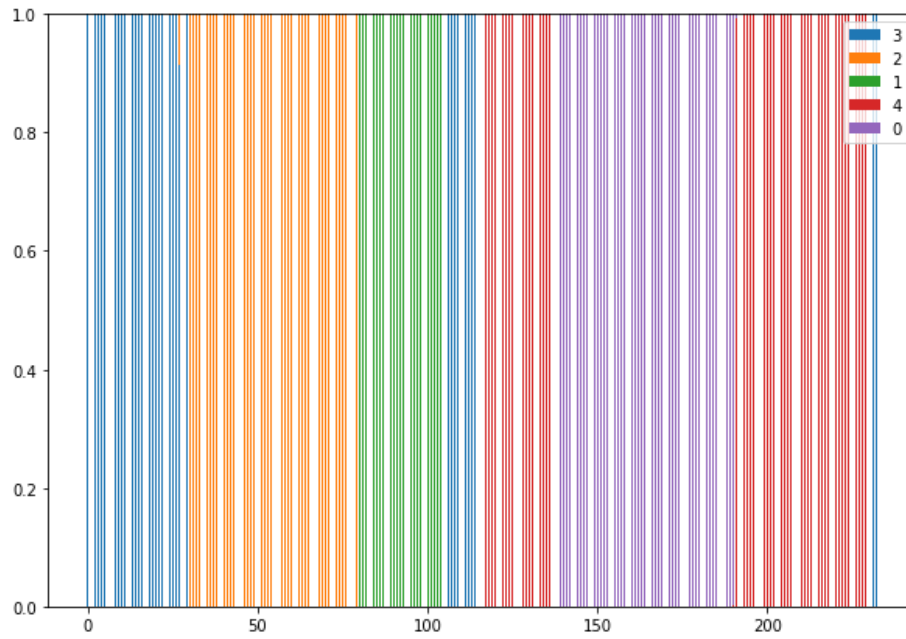


Figura 12: Evolución de las Probabilidades GMM.

De la figura 12 es posible visualizar que la asignación de probabilidades de pertenencia a cada cluster a lo largo del periodo es casi absoluta, implicando que el algoritmo al optimizar es capaz de capturar la diferencia en los datos con los parámetros entregados (al igual que en KM). Por otro lado, es interesante el hecho de que si se compara la evolución de las probabilidades entre KM y GMM, se encuentra que los periodos que evidencian son parecidos y, la principal diferencia viene de que el último periodo de KM presenta un periodo extra en GMM.

Por último, hay que recalcar que esta técnica es una propuesta alternativa a lo que propone la publicación y, por consiguiente, se presenta con el objetivo de agregar valor a la discusión. Esto debido a la relevancia de la probabilidad de pertenencia a cada fase del ciclo económico en los resultados del modelo.

#### 8.2.4. Cálculo y Optimización del Portafolio final

El cálculo del portafolio final se realiza tomando en consideración los retornos esperados por fase para cada clase de activo, la matriz de varianza-covarianza de los retornos y la función objetivo de maximizar el retorno esperado, todo esto sujeto a una volatilidad objetivo y restringiendo el tamaño de las posiciones a las que puede tomar la AFP.

## Retorno Esperado por Fase

Para estimar los retornos esperados por fase se utilizan regresiones lineales, donde la variable dependiente es el retorno esperado a 6 meses de las clases de activos y las variables independientes son las probabilidades de estar en cada fase, por lo tanto, los B's asociados a estas probabilidades son los retornos esperados del activo para la fase. Se utilizan los retornos esperados a 6 meses por dos razones, en primer lugar, porque es un tiempo prudente para que los cambios en las variables macroeconómicas repercutan en el mercado y, en segundo lugar, es el horizonte de tiempo que usa el mánager.

Se utilizan regresiones porque al ser un método de aprendizaje supervisado es posible generar una función que relacione las probabilidades asociadas a una fase con los retornos y, de esta manera, caracterizar los retornos esperados por fase<sup>55</sup>.

Se tiene que el retorno esperado a 6 meses para un activo  $i$  es:

$$RE_{i,t} = \frac{P_{i,t+6}}{P_{i,t}}$$

La regresión utilizada fue:

$$RE_{i,t} = B_1 * ProbFase1_{i,t} + B_2 * ProbFase2_{i,t} + B_3 * ProbFase3_{i,t} + B_4 * ProbFase4_{i,t}$$

Donde:

- $B_i$ : Retorno esperado para el activo en la fase  $i$

Se impone el intercepto igual a 0 debido a que no tiene sentido económico en caso contrario, puesto que implicaría que existe un retorno esperado base. Utilizando toda la muestra y realizando la regresión para cada clase de activo se obtienen los siguientes resultados para el 31 de mayo de 2019:

---

<sup>55</sup> También se diseñó una técnica de  $k$  vecinos más cercanos para determinar los retornos esperados por fase, sin embargo, carecía de criterio objetivo, puesto que asume que en un periodo  $t$  se estaba en fase  $j$  cuando la probabilidad de  $j$  era mayor al resto y, además, los resultados fueron muy malos. Por lo tanto, no agrega valor a la discusión incluirlo.



	0	1	2	3
.DMEQLCEW Index	7.54%	-9.53%	13.46%	-17.03%
.EMEQCULC Index	1.80%	2.00%	16.88%	-4.20%
.CHFICUAA Index	2.99%	5.03%	4.31%	4.04%
LUATTRUU Index	0.26%	5.02%	0.74%	5.21%
.HBFIWOCO Index	0.60%	16.84%	5.23%	-5.79%

Tabla 9: Retornos Esperados por Fase para Cada Activo. KM.

	0	1	2	3	4
.DMEQLCEW Index	5.37%	-11.66%	6.61%	-5.05%	2.02%
.EMEQCULC Index	0.57%	-0.22%	11.21%	3.57%	3.81%
.CHFICUAA Index	3.86%	3.98%	4.07%	4.50%	3.34%
LUATTRUU Index	1.19%	4.51%	1.71%	3.94%	1.51%
.HBFIWOCO Index	2.29%	-1.37%	5.09%	10.29%	3.42%

Tabla 10: Retornos Esperados por Fase para Cada Activo. GMM.

### Matriz de Varianza Covarianza

Para determinar la matriz de varianza-covarianza para el periodo T se utiliza la siguiente formula. Sea A y B dos vectores de retorno esperado de diferentes clases de activos y  $REX_i$  el retorno esperado para el activo X en la fase i:

$$\sigma_{A,B}^T = \frac{\sum_{t \leq (T-6)} \left[ \left( A_t - \sum_{0 \leq i \leq 3} \frac{P_i * REA_i}{T} \right) * \left( B_t - \sum_{0 \leq i \leq 3} \frac{P_i * REB_i}{T} \right) \right]}{T - 1}$$

El objetivo de calcularla así viene de que la matriz se ajusta a las probabilidades de estar en cada fase del ciclo en el periodo a calcular, capturando de cierta forma las relaciones de riesgo en el tiempo estudiado. Además, es relevante destacar que los retornos esperados utilizados son hasta T-6, de esta manera no se está utilizando información del futuro.

Por otro lado, tal y como exponen Bourgeron, Lezimi, Roncalli (2018) el problema es sumamente sensible a esta matriz de varianza-covarianza, por lo que es recomendable realizarle ajustes. También está lo expuesto por Husnain, Hassan, Lamarque (2016) que muestran empíricamente que los métodos más sofisticados de ajuste no son significativamente mejores que los más básicos, además, postulan que un buen ajuste es:  $VarCovar = 0.5 * VarCovar + 0.5 VarCovar_{diagonal}$ , es decir, modifican las covarianzas entre los activos reduciéndolas a la mitad.

A su vez, López de Prado (2019) expone la inestabilidad de los resultados de la optimización dadas las características matemáticas del problema que se resuelve. Considerando que se tiene una volatilidad objetivo como restricción, al momento de optimizar se utiliza el inverso de la matriz de varianza-covarianza, con esto en mente expone lo siguiente:

Sea  $X$  una matriz de observaciones aleatorias de tamaño  $T \times M$  de media 0 y varianza  $\sigma^2$ , la matriz  $C = T^{-1} * X^T * X$  tiene valores propios  $\lambda$  que convergen a la función distribución de probabilidad Marchenko-Pasteur<sup>56</sup>:

$$f[\lambda] = \begin{cases} \frac{T \sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{N 2\pi\lambda\sigma^2} & \text{if } \lambda \in [\lambda_-, \lambda_+] \\ 0 & \text{if } \lambda \notin [\lambda_-, \lambda_+] \end{cases}$$

Con lo siguientes valores propios máximos y mínimos esperados:

$$\lambda_+ = \sigma^2 \left(1 + \sqrt{\frac{N}{T}}\right)^2, \lambda_- = \sigma^2 \left(1 - \sqrt{\frac{N}{T}}\right)^2.$$

Con esto en consideración, cuando  $\sigma^2 = 1$ , se tiene que  $C$  es la matriz de correlaciones de  $X$ .

Por otro lado, dadas las características del problema, se tiene que si se toma como  $X$  a la matriz  $VarCovar$  y, asumiendo que el promedio es 0 y  $\sigma^2 = 1$ , se tiene que  $\frac{N}{T} \rightarrow 1$ , de esta manera,  $\lambda_- \rightarrow 0$ , con lo que el determinante de  $VarCovar$  sería aproximadamente 0, implicando que  $VarCovar^{-1}$  no pueda ser estimado robustamente.

Para resolver este problema, López de Prado (2019) propone calcular los vectores y valores propios asociados a la matriz de correlaciones calculada a partir de la matriz de varianza-covarianza y, luego, ajustar la distribución Marchenko-Pastur a la distribución empírica de los valores propios. Con esto es posible separar los valores propios relacionados con el ruido con los relacionados a la señal de las clases de activos, para luego modificar los relacionados con el ruido y reconstruir la matriz de varianza-covarianza.

<sup>56</sup> Wikipedia. Distribución Marchenko-Pastur. Disponible en: [https://en.wikipedia.org/wiki/Marchenko%E2%80%93Pastur\\_distribution](https://en.wikipedia.org/wiki/Marchenko%E2%80%93Pastur_distribution) (consulta: 10 de enero de 2020).

También, se propone realizar los siguientes ajustes a la matriz de varianza-covarianza:

- Rolling: Corresponde a tomar un periodo de tiempo comprendido entre  $t' > 0$  hasta  $T$  para calcular la matriz de varianza-covarianza, por ejemplo, rolling 48 meses implica tomar los últimos 48 datos más recientes para calcular la matriz. El objetivo de este método es utilizar una ventana de tiempo móvil que a medida que se vaya incorporando nueva data se vaya actualizando y avanzando, esto para considerar las relaciones entre las variables para periodos más actuales, por ejemplo, si se está en 2018 y en 2006 (12 años atrás) hubo una crisis financiera, al considerar toda la data desde el año 2000 en adelante se agrega mayor volatilidad a los activos más riesgosos como renta variable, implicando que el allocation en estos activos sea menor.
- EWMA: Corresponde a una técnica de series de tiempo la cual utiliza un factor  $\lambda$  para ponderar la varianza-covarianza histórica y un factor  $(1 - \lambda)$  para ponderar el error (la diferencia entre el valor real y el pronosticado, el cual se puede interpretar como “noticias”). Por lo tanto, tiene como objetivo ajustar, en cierta medida, las variaciones que pasan en el último periodo. Lo anterior se puede traducir de la siguiente manera:

$$\sigma_{A,B}^T = (1 - \lambda)(\varepsilon_A^{T-1})(\varepsilon_B^{T-1}) + \lambda\sigma_{A,B}^{T-1}$$

Donde:

- $\varepsilon_X^T = X_T - \sum \frac{\sum_{0 \leq i \leq 3} P_i * REX_i}{T}$
- $\sigma_{A,B}^1 = \frac{1}{T-1} \sum_{t \leq T-1} \varepsilon_A^t * \varepsilon_B^t$

Es importante considerar que cuando se utiliza EWMA en el modelo se calcula la matriz de varianza-covarianza para el periodo  $T+1$ , es decir, se tiene como resultado el pronostico de la matriz para el 30 de junio de 2019 considerando la data hasta el 31 de mayo de 2019. De esta manera se genera allocation considerando la matriz de varianza-covarianza esperada. Por último, J.P. Morgan (1993) sugiere un  $\lambda = 0.94$

## Optimización

Se considera el siguiente problema de optimización (Markowitz), donde  $\mu^T$  es el vector de retornos esperado de los asset class y  $x_i$  es la posición asociada al activo i, con  $\bar{x} = (x_1, x_2, x_3, x_4, x_5)$  (RVD, RVE, RFL, RFIG, RFHY):

$$\begin{aligned} & \text{s.t.} && \text{Max } \mu^T \bar{x} \\ & && \bar{x}^T * \text{VarCovar} * \bar{x} \leq \text{TargetVol}^2 \\ & && \text{limite}_{inferior}^i \leq x_i \leq \text{limite}_{superior}^i, \forall i \in (1,5) \\ & && \text{limite}_{inferior}^j \leq \sum x_j \leq \text{limite}_{superior}^j, \quad j \in \text{renta variable} \\ & && \text{limite}_{inferior}^k \leq \sum x_k \leq \text{limite}_{superior}^k, \quad k \in \text{renta fija} \end{aligned}$$

Además, se consideran dos casos posibles para generar el allocation:

- No considerar cash para cerrar la posición, implicando la siguiente restricción:

$$1^T \bar{x} = 0$$

- Considerar cash para cerrar la posición, implicando la siguiente restricción:

$$\text{limite}_{inferior}^{cash} \leq 1^T \bar{x} \leq \text{limite}_{superior}^{cash}$$

El modelo maximiza el retorno esperado restringido a una volatilidad objetivo, la cual es entregada por el mánager a cargo. Esta es la volatilidad que espera la compañía al hacer allocation mes a mes. Por otro lado, se restringe el tamaño de la posición que se puede colocar a nivel de asset class en particular, como también a nivel de tipo de asset class (renta variable y renta fija). Por último, como el problema se trata de una optimización con restricción cuadrática, se utiliza la técnica de penalización para encontrar el resultado óptimo.

A su vez, se resuelve el siguiente problema de optimización (ranking):

$$\text{Max } \mu^T (\text{pond} * x)$$

s.a:

$$x \geq 0$$

$$(\text{pond} * x)^T * \text{VarCovar} * (\text{pond} * x) \leq \text{TargetVol}^2$$

$$x \leq \frac{\text{limite}_{\text{posición}}}{2}$$

El modelo optimiza el retorno esperado considerando un ponderador  $\text{pond} = (2,1,0,-1,-2)$ , el cual obliga a tomar posiciones grandes en los activos que se espera tengan un rendimiento superior y posiciones pequeñas en activos que se espera tengan un rendimiento inferior, por lo que “x” es el factor por el que se pondera para que se cumpla la restricción de volatilidad.

Los límites se establecen en base a las posiciones posibles que puede tener la compañía, en relación con los costos asociados y a la gestión de los activos para ejecutar los movimientos.

También es importante mencionar que como se tienen retornos esperados distintos para cada fase del ciclo, el problema se resuelve 4 veces para cada modelo de optimización (una por cada fase del ciclo), implicando que se genera un portafolio óptimo según fase considerando los distintos  $\mu^T$  y, a su vez, estos son ponderados según la probabilidad de pertenecer a cada ciclo, por ejemplo:

31-05-19	Prob				Final
	43%	12%	23%	22%	
Fase	0	1	2	3	
R.V.D.	-0.111%	-0.075%	-0.039%	0.008%	-0.063%
R.V.E.	0.018%	0.072%	0.040%	-0.030%	0.019%
R.F.L.	0.187%	0.109%	0.319%	0.336%	0.241%
R.F.I.G.	-0.179%	-0.034%	-0.344%	-0.341%	-0.235%
R.F.H.Y.	0.085%	-0.072%	0.025%	0.027%	0.039%

Tabla 11: Ejemplo de Allocation Final.

## 8.2.5. Análisis y Backtest

Es importante destacar el hecho de que, dada la diferencia entre los métodos para calcular las probabilidades (KM y GMM), donde la cantidad de fases y las probabilidades son distintas, es necesario realizar el análisis de cada técnica como modelos independientes, aún cuando se utilicen las mismas técnicas para calcular los retornos esperados por fase y la matriz de varianza-covarianza.

### Backtest

A continuación, se muestra el backtest para distintas estrategias de los modelos KM y GMM. Este, en primera instancia, considera un periodo desde el 31 de enero de 2013 hasta el 31 de mayo de 2019. Se determina que el 31 de enero de 2013 es una buena fecha de inicio ya que, tal y como se puede apreciar en la evolución de las probabilidades de KM y GMM (figura 10 y 12), desde el 31 de enero de 2000 hasta el 31 de enero de 2013 se tienen las 4 fases de KM y las 5 fases de GMM identificadas, con lo que se pueden entrenar los modelos de buena manera.

A su vez, se determinan como límite superior 1% y como límite inferior -1% tanto a nivel de activo por sí solo como de la suma por tipo de activo (renta variable y renta fija). La lógica que sigue el backtest es la siguiente: A medida que se avanza en el tiempo se agrega la nueva data y, con la nueva muestra, se recalculan las probabilidades<sup>57</sup> y se mapea de mejor manera (con más datos) la función que relaciona el retorno esperado con las probabilidades para generar la optimización.

Para construir el backtest se calculan los retornos mensuales con el valor de los precios de los índices de cada clase de activo desde el 31 de enero de 2013 hasta el 31 de mayo de 2019, luego, estos retornos son ponderados por la posición correspondiente para finalmente ser sumados y, así, determinar el retorno de la estrategia.

Por otro lado, en un inicio se calcula el retorno acumulado para cada estrategia con las siguientes matrices de varianza-covarianza, que se identifican como buen punto de inicio para verificar la performance del modelo en sí, y no como combinaciones que causen mejoras (por ejemplo, Roll-Ewma):

- Var-Covar Data Completa: Utiliza toda la data que se tiene hasta la fecha en la que se hace allocation para calcular la matriz de varianza-covarianza de los retornos (31 de enero del 2000 hasta el 31 de enero de 2013 para el primer allocation).

---

<sup>57</sup> Es importante destacar que, si bien las probabilidades se van actualizando, estas a medida que se va agregando no cambian en demasía, implicando que se identifican los mismos periodos a través del tiempo. Se adjunta en anexo 4 la evolución de las probabilidades.

- Ewma: Utiliza EWMA considerando toda la data que se tenga hasta la fecha en que se hace el allocation (31 de enero del 2000 hasta el 31 de enero de 2013 para el primer allocation).
- Rolling 48M: Utiliza los últimos cuatro años de data para calcular la matriz de varianza-covarianza de los retornos.
- Rolling 36M: Utiliza los últimos tres años de data para calcular la matriz de varianza-covarianza de los retornos.
- Rolling 24M: Utiliza los últimos dos años de data para calcular la matriz de varianza-covarianza de los retornos.

### Simulaciones

Se propone realizar simulaciones dado que el periodo a estudiar para cada uno de los modelos está muy concentrado en una o dos fases, por lo tanto, es interesante analizar cuál es la performance en otra fase y, más aún, ver como distribuyen los retornos de las estrategias en distintos periodos y no en uno en particular.

El modelo de simulaciones funciona de la siguiente manera, en primer lugar, se elige la cantidad de simulaciones a realizar, luego se escoge una fecha de inicio aleatoria, para continuar con la selección de N datos desde esa fecha para determinar la performance de la estrategia, posterior a esto tanto la data macro como los retornos de los índices de las fechas seleccionadas son trasladadas al final de la base de datos, implicando que el modelo sea entrenado con toda la data menos la que se va a testear, manteniendo así la característica de entrenar los modelos con todas las fases, para finalmente calcular la performance del periodo estudiado.

Se determina N igual a 12 porque, como se puede observar en la evolución de las probabilidades de KM y GMM (figuras 10 y 12), en general, las fases tienen un largo aproximado de 2 años, por lo que tomar un periodo mayor a este implica que en algunos casos el modelo se entrena sin existir una fase, sin embargo, a juicio de experto y dado que los modelos tienen características a largo plazo (dado que utilizan data macro), este periodo puede parecer relativamente corto, puesto que si bien no hay una medida estándar para el largo plaza, 1 año se considera como el tiempo mínimo.

## 8.2.6. K-means

### Caracterización de las Fases

En base a los resultados obtenidos de la regresión donde se estimaron los rendimientos esperados para cada clase de activo en cada fase y, por otro lado, lo expuesto por Defend, Portelli y Cesarini (2017) donde plantean que:

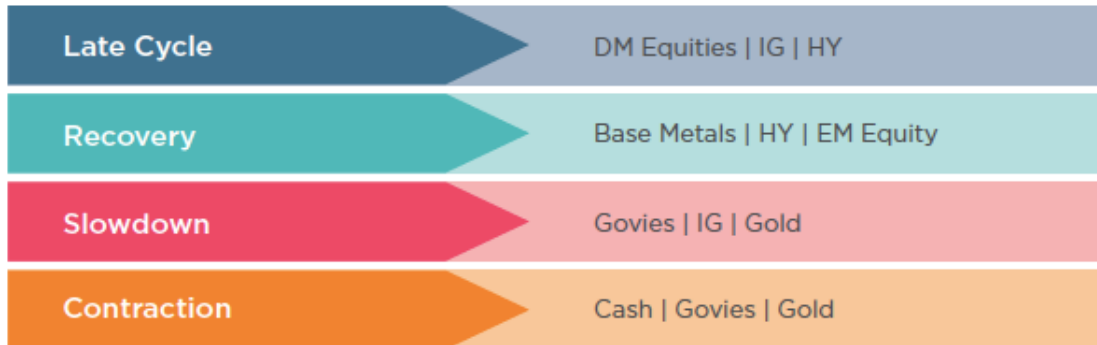


Figura 13: Características de los Ciclos.

Se concluye que:

- Fase 0: Expansión (late cycle).
- Fase 1: Slowdown.
- Fase 2: Recuperación.
- Fase 3: Contracción (recesión).

Tal como se muestra a continuación:

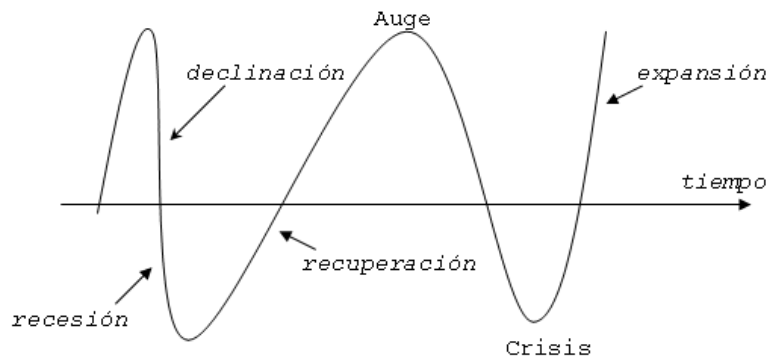


Figura 14: Características de los Ciclos<sup>58</sup>.

<sup>58</sup> Inma Economía. Ciclos Económicos. Disponible en: [http://inma-economia2.blogspot.com/2014/06/los-ciclos-economicos-y-la-crisis\\_2.html](http://inma-economia2.blogspot.com/2014/06/los-ciclos-economicos-y-la-crisis_2.html) (consulta: 30 de agosto de 2019).



## Backtest

Para construir el backtest se utiliza la lógica expuesta en un inicio, la que se resume a continuación:

- Periodo: 31 de enero 2013 hasta el 31 de mayo de 2019.
- Límites superiores e inferiores de  $\pm 1\%$  tanto por clase de activo como por suma de tipo de activo (renta fija y renta variable).
- Volatilidad objetivo de 0.02% (2 puntos base).

Fase del periodo a backtestear:

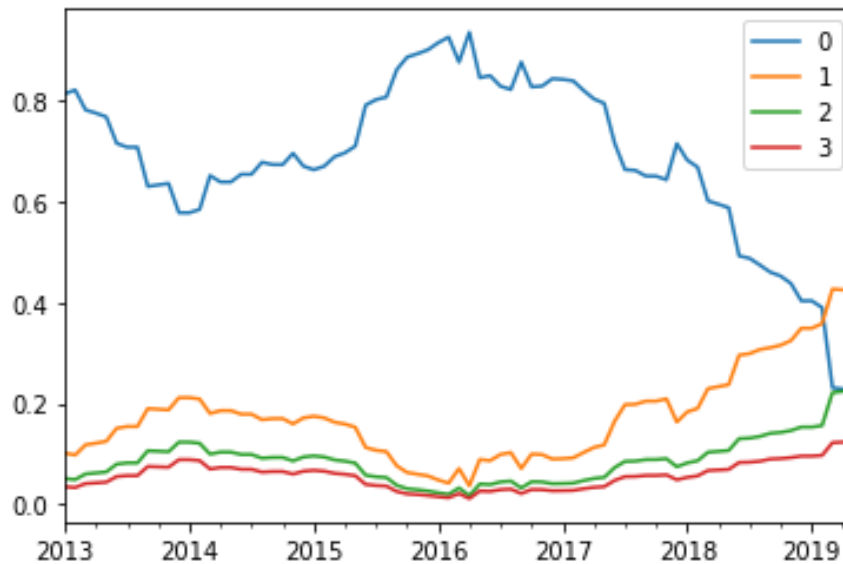


Figura 15: Fases Backtest.

Como es posible evidenciar, el periodo a estudiar se encuentra gran parte del tiempo en expansión para luego cambiar a slowdown. A continuación, se muestran los retornos esperados para cada una de estas fases el 31 de mayo de 2019, los cuales deberían poseer el mismo signo para todo el intervalo a estudiar:

	<b>Expansión</b>	<b>Slowdown</b>
<b>.DMEQLCEW Index</b>	7.54%	-9.53%
<b>.EMEQCULC Index</b>	1.80%	2.00%
<b>.CHFICUAA Index</b>	2.99%	5.03%
<b>LUATTRUU Index</b>	0.26%	5.02%
<b>.HBFIWOCO Index</b>	0.60%	16.84%

Tabla 12: Retornos Esperados Fases Backtest.

## Markowitz:

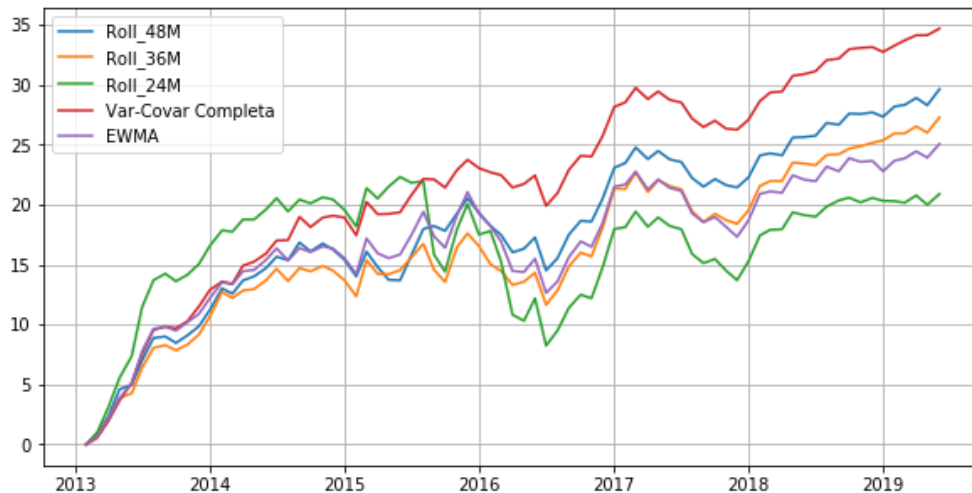


Figura 16: Retornos Acumulados por las Estrategias. [PB]

De la imagen es posible visualizar lo siguiente. En primer lugar, mayor es el retorno obtenido a medida que más data se utiliza para calcular la matriz de varianza-covarianza, en segundo lugar, la forma de las curvas es bastante similar y, en tercer lugar, el retorno acumulado máximo es de 0,35% (35 PB) en un periodo de 6 años, el cual no es sobresaliente considerando que con una volatilidad objetivo de 2 PB mensual se espera en un caso excepcional un retorno de 1,44% (que todos los meses se obtenga el 0.02%<sup>59</sup>) y en un caso bueno un retorno de 0.44% (que todos los años se obtenga la volatilidad anual<sup>60</sup>). Por lo tanto, se concluye que las distintas estrategias que aparecen en la figura 16 toman posiciones en su mayoría en renta fija, implicando buena performance y bajo retorno, esto también explicaría que al tomar menos data (o ponderar más la última data) se disminuya la rentabilidad, puesto que se toman mayores posiciones en instrumentos de renta variable (que al parecer no tuvieron buen performance), ya que es lógico pensar que la varianza de estos disminuya.

Por otro lado, se evidencia que desde al inicio hasta la mitad del 2016 hubo perdidas sistemáticas, especialmente en las estrategias más volátiles (que utilizan menos data para calcular la matriz de varianza-covarianza), lo que implica alguna variación importante en los mercados de renta variable. Como antecedente se tiene que la moneda China (presente en el índice de renta variable emergente) fue devaluada luego de presentar una fuerte revaluación<sup>61</sup> para relanzar su economía en esos momentos debilitada.

<sup>59</sup> Es decir,  $r_a = 0.0002 * x$ , con x: cantidad de meses.

<sup>60</sup> Es decir,  $r = 0.0002 * \sqrt{12} * x$ , con x: cantidad de años.

<sup>61</sup> El Financiero. China Entreta en 2016 Mayor Presión por su Moneda. Disponible en: <https://www.elfinanciero.com.mx/mercados/divisas/china-enfrentara-en-2016-mayor-presion-por-su-moneda-el-yuan> (consulta: 17 de octubre de 2019).

También se analiza la opción de cerrar las posiciones con caja CLP para ver la posibilidad de tomar más riesgo, imponiendo que el límite de posición en caja CLP es de  $\pm 1\%$ . Obteniéndose los siguientes resultados:

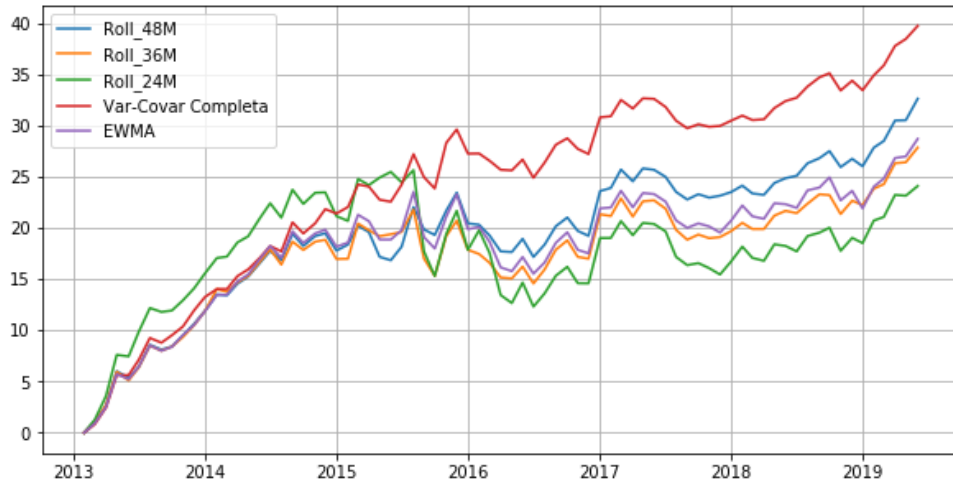


Figura 17: Rentabilidad Acumulada Estrategias con Cash. [PB]

Las conclusiones no difieren de las ya expuestas. La performance mejora en un bajo porcentaje, por lo que es posible razonar que se toman posiciones más grandes en activos de renta fija.

Lo anterior se puede verificar comparando los portafolios de las distintas estrategias con la opción de considerar o no caja chilena para cerrar las posiciones, obteniéndose los siguientes resultados para el Rolling 24M (el que se espera tome mayores posiciones en renta variable):

Index	DMEQLCEW Inde:	EMEQCULC Inde>	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	18.5465	-18.2366	71.7946	-77.4343	5.32981
std	23.4555	24.1667	17.8197	20.3714	28.0294
min	-34.0733	-60.6683	14.0208	-95.1555	-55.5048
25%	3.42929	-34.9063	60.7925	-90.2247	-17.9331
50%	15.8331	-26.4915	71.8535	-86.8095	4.50348
75%	31.1143	2.54053	87.87	-75.18	27.4638
max	66.656	42.137	94.4394	-21.6568	59.058

Figura 18: Caracterización Portafolio Rolling 24M sin Caja CLP. [PB]

Index	DMEQLCEW Inde	EMEQCULC Inde	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index	LVACLI Index
count	77	77	77	77	77	77
mean	32.633	-21.4253	94.9681	-33.3539	25.6413	-98.4632
std	21.0395	29.8822	8.31544	31.3785	35.2619	2.46212
min	-15.2264	-66.7965	60.8709	-87.2768	-42.2169	-100
25%	19.8361	-41.4248	93.5602	-52.0504	-8.19141	-100
50%	30.7791	-35.5448	99.7813	-29.5633	25.8186	-99.5569
75%	43.1317	0.686659	100	-19.9471	49.2152	-97.424
max	80.5382	40.7599	100	61.1133	83.9838	-83.307

Figura 19: Caracterización Portafolio Rolling 24M con Caja CLP. [PB]

Es posible evidenciar que en ambos casos las posiciones más pequeñas (en términos de tamaño) se encuentran, en general, en renta variable. Si bien el agregar la opción de cerrar la posición con caja chilena implica, en promedio, un aumento del tamaño de las posiciones en renta variable emergente, también lleva a tope la posición en renta fija local, la cual en la opción sin caja ya es grande y siempre positiva en ambos casos. Entonces, es posible relacionar la buena performance de las estrategias con la renta fija local, lo que a su vez explicaría el bajo retorno acumulado en los 6 años que se analiza la estrategia.

A su vez, de toda la clase de activos en los que se puede invertir, se tiene que renta fija local es el mercado más pequeño y, por lo tanto, el menos líquido. Entonces, es razonable disminuir los límites de inversión de este, según juicio de experto, a un 0.25%. Se obtienen los siguientes resultados:

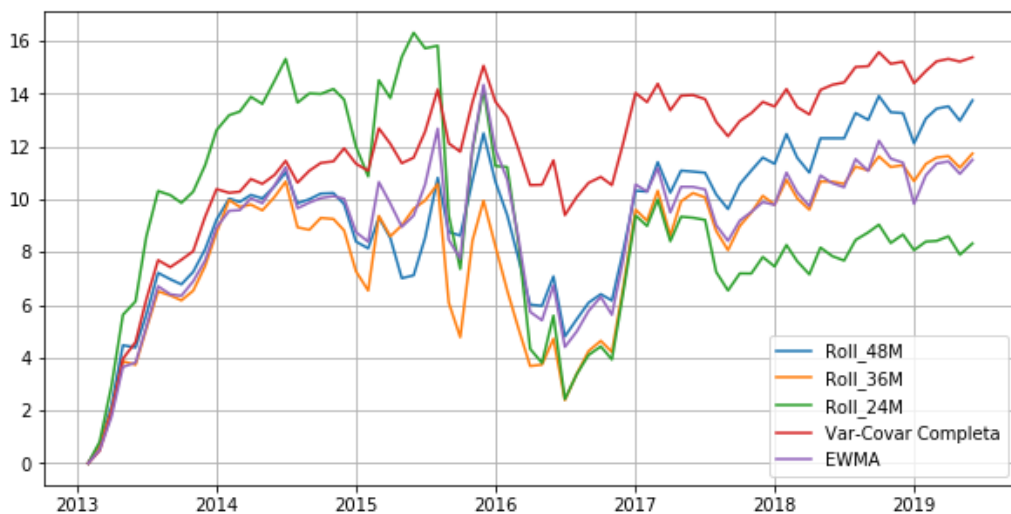


Figura 20: Rentabilidad Acumulada Estrategias Markowitz RFL Restringida. [PB]

Tal y como se observa, la performance baja considerablemente en cada una de las estrategias, lo que confirma el hecho de que la ganancia se genera en su mayoría por renta fija y, en específico, renta fija local, la que a su vez es bastante ilíquida, lo que implica que la estrategia no es posible de implementar. Más aún si se considera el hecho de que, a juicio de experto, no es atractivo un portafolio centrado netamente en renta fija puesto que el exceso de retorno que se puede generar es considerablemente menor al que se puede lograr utilizando renta variable.

Ranking:

Considerando las mismas condiciones iniciales con respecto a los límites expuestas en el caso inicial anterior, se obtienen los siguientes resultados:



Figura 21: Rentabilidad Acumulada Estrategias Ranking. [PB]

De la figura es posible evidenciar que, en primer lugar, también se tiene el efecto del 2016 en la performance, en segundo lugar, la performance es menor a la mitad del caso anterior, una posible razón es que la regresión lineal no genere una buena función que caracterice los retornos esperados por fase, otra posible razón es que considerando que esta formulación obliga a tomar posiciones grandes donde se espera obtener mejores resultados y chicas es caso contrario, puede ocurrir que el tamaño de las posiciones grandes per se sean pequeños (dada la restricción de volatilidad). Para evidenciar esto se muestran las características del portafolio rolling 36 meses que tiene el mejor performance y rolling 24 meses que se espera sea el que tome posiciones más grandes en renta variable:

Index	DMEQLCEW Index	EMEQCULC Index	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	25.7509	-24.0007	17.6024	-20.8038	1.45112
std	12.6719	17.5351	5.15843	12.3808	9.66496
min	-10.9031	-52.2986	2.30126	-50.109	-22.4834
25%	20.6274	-37.3485	15.8752	-26.348	-6.23928
50%	28.0826	-30.1843	17.9458	-19.1483	4.56556
75%	32.1599	-1.38112	19.9005	-15.659	5.72625
max	46.858	0.203597	25.7128	3.0901	24.2543

Figura 22: Caracterización Portafolios Ranking Rolling 36 Meses. [PB]

Index	DMEQLCEW Index	EMEQCULC Index	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	29.2219	-27.1892	19.6094	-21.7375	0.0953959
std	18.372	22.6395	8.44774	14.7281	13.4099
min	-8.46152	-79.1212	2.16068	-83.1334	-34.8374
25%	17.5508	-43.1669	15.1299	-27.9162	-5.70301
50%	23.5248	-30.3029	17.8015	-20.2723	4.73614
75%	37.8171	-2.23901	22.9678	-14.1463	6.05411
max	79.8153	0.197759	42.353	3.71883	28.3306

Figura 23: Caracterización Portafolios Ranking Rolling 24 Meses. [PB]

De las figuras es posible evidenciar que, en general, se toman posiciones más pequeñas en renta fija que en el caso anterior, resultado esperado debido a la caracterización de gran parte de la fase que se presenta en el periodo a estudiar, sin embargo, las características de las posiciones per se son similares, por ejemplo, RFL está siempre largo y es en la que más se invierte en términos de tamaño.

Dado la estructura de la formulación de este problema, no es posible cerrar posiciones con caja chilena, sin embargo, al igual que en el caso anterior de Markowitz, no se esperan mejoras sustanciales, a su vez, y dada las características de las posiciones de la estrategia rolling 24 y 36 meses, no es esperable que al reducir la posición en renta fija local se obtengan mejores resultados.

Dado los retornos esperados en cada fase para cada clase de activo, se desprende que los resultados de ambos modelos están sujetos en gran medida a la matriz de varianza-covarianza que se utiliza para cumplir con la restricción de volatilidad. Por lo tanto, uno de los primeros pasos a realizar es analizar la performance del modelo sin volatilidad objetivo, lo que se muestra en el análisis de sensibilidad. Por otro lado, no se analizan combinaciones del tipo Rolling-EWMA, Rolling-EWMA-caja, etc. puesto que tal y como se muestra en el análisis, no se espera que el rendimiento mejore sustancialmente, lo que es el objetivo del trabajo.

Por último, es posible apreciar que la performance de todas las estrategias de 2013 a 2014 evidencia, en general, una pendiente positiva considerable y, por el contrario, post ese periodo hubo una disminución considerable del rendimiento. Dicho esto, es posible razonar que luego de ese periodo ocurrieron acontecimientos que no existieron antes, implicando que las relaciones encontradas en un inicio entre las probabilidades y el retorno no se ajustaran de buena manera a las condiciones que en ese periodo se presentan, o también, la existencia de una quinta fase.

### **8.2.7. Análisis de sensibilidad – KM**

#### Modelo Irrestringido en Volatilidad

Dada las características del modelo, donde en base a la matriz de varianza-covarianza se ajustan las posiciones, resulta interesante analizar la performance de las estrategias cuando el modelo es irrestringido en volatilidad, es decir, sin considerar una volatilidad objetivo, esto implica que toma posiciones considerando el mejor retorno esperado, a su vez, la restricción de volatilidad viene implícita de la restricción en el tamaño de las posiciones (que es una medida menos subjetiva y sensible, puesto que es determinada por el mánager en base a la cantidad de dinero que maneja y puede mover la compañía), tanto por clase de activo como por tipo (renta fija y renta variable), por lo tanto, resulta sugerente determinar el tope de volatilidad que viene determinado por esta restricción. Considerando el caso base que toma como límite superior 1% y límite inferior -1% tanto como activo por si solo y la suma de tipos de activos (renta fija y renta variable):

Markowitz:



Figura 24: Rentabilidad Acumulada Modelo sin Volatilidad Objetivo. [PB]

Lo primero que llama la atención es que se aprecia mayor volatilidad en la performance después del año 2016 y, como es de común saber, la renta variable es más volátil que la renta fija, por lo tanto, se subentiende que se toman mayores posiciones en estos activos. Por otro lado, resulta interesante analizar la volatilidad promedio de las estrategias, obteniéndose para var-covar completa (más aversión al riesgo) y Rolling 24 meses (menos aversión al riesgo) los siguientes resultados:

<b>Volatilidad [BP]</b>	<b>Roll 29 M</b>	<b>Var-Covar Completa</b>	<b>Promedio</b>
<b>Promedio</b>	1.84	2.55	2.195
<b>Mediana</b>	1.91	2.71	2.31
<b>Minimo</b>	0.41	0.93	0.67
<b>Maximo</b>	3.11	4.12	3.615

Tabla 13: Volatilidad Irrestringido. [PB]

Al no estar restringido por volatilidad objetivo y con los resultados obtenidos, es posible concluir que, en general, los límites impuestos al modelo en este periodo estudiado implican aversión al riesgo, puesto que se tienen volatilidades bajas considerando tanto la matriz con mayor y menor aversión al riesgo, además, se tiene que considerar que los resultados están sujetos a los retornos esperados por fase, y estos pueden ser la mayor piedra de tope para obtener mejores resultados, puesto que si no son caracterizados de buena manera no se obtienen resultados sobresalientes.



Al comparar la estrategia sin volatilidad objetivo con respecto al resto con dos puntos bases, se evidencio lo siguiente:

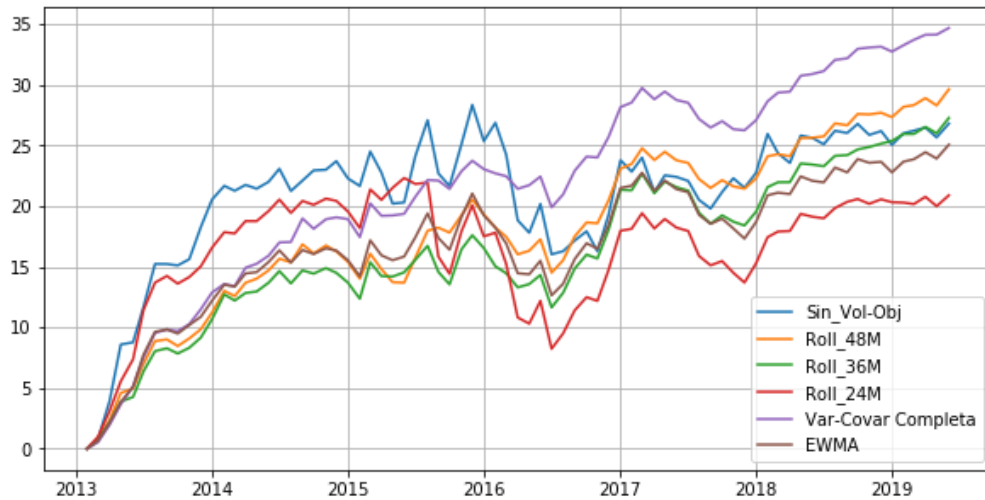


Figura 25: Rendimiento Estrategias TV 2 BP vs sin TV. [PB]

En este caso particular, es posible evidenciar que mayor riesgo en algunas posiciones no signifioco mayor retorno, por lo que las estrategias utilizadas no son optimizas en la generación del allocation y poseen un espacio de mejora. También, es relevante destacar el hecho de que las estrategias, en su mayoría, acumulan todo su performance entre los periodos correspondidos de 2013 y mediados de 2014, y mediados del 2017 hasta el 31 de mayo de 2019, esto porque el periodo entre mediados de 2014 y mediados del 2017 se considera como “flat”, es decir, que no genera retorno.

Ranking:



Figura 26: Rentabilidad Ranking Sin Volatilidad Objetivo. [PB]

Como la formulación del problema toma las mismas posiciones, pero ahora con un mayor peso dado que no se está restringido en volatilidad, implicando que se amplifique la forma de las curvas mostradas en el apartado anterior. Dada la restricción de los límites, es posible tomar las posiciones máximas que estos establecen. Esto se puede apreciar en la siguiente imagen, donde se contrastan las curvas con volatilidad objetivo de 2 puntos base con sin volatilidad objetivo:

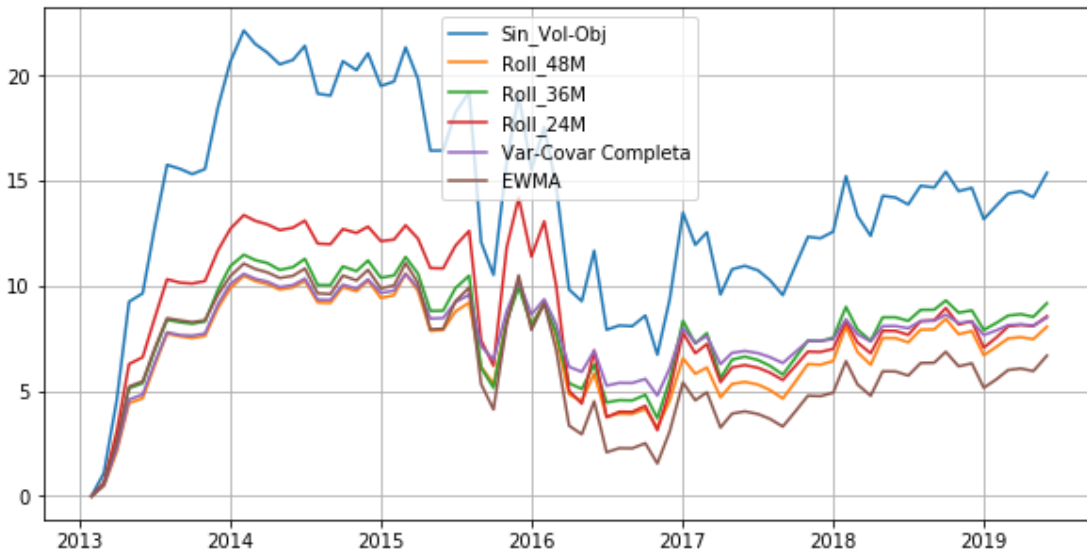


Figura 27: Rendimiento TV 2 BP vs Sin Volatilidad Objetivo Estrategia Ranking. [PB]

### Agregar Una Nueva Fase

El agregar una nueva fase tiene como objetivo generar valor diferenciando la data aún más, por lo que, en caso de existir, el comportamiento de la data estaría mejor definido implicando que se tome una mejor decisión. Al comparar los resultados de las estrategias con mejor performance (data complete y roll 48 meses) se obtiene:

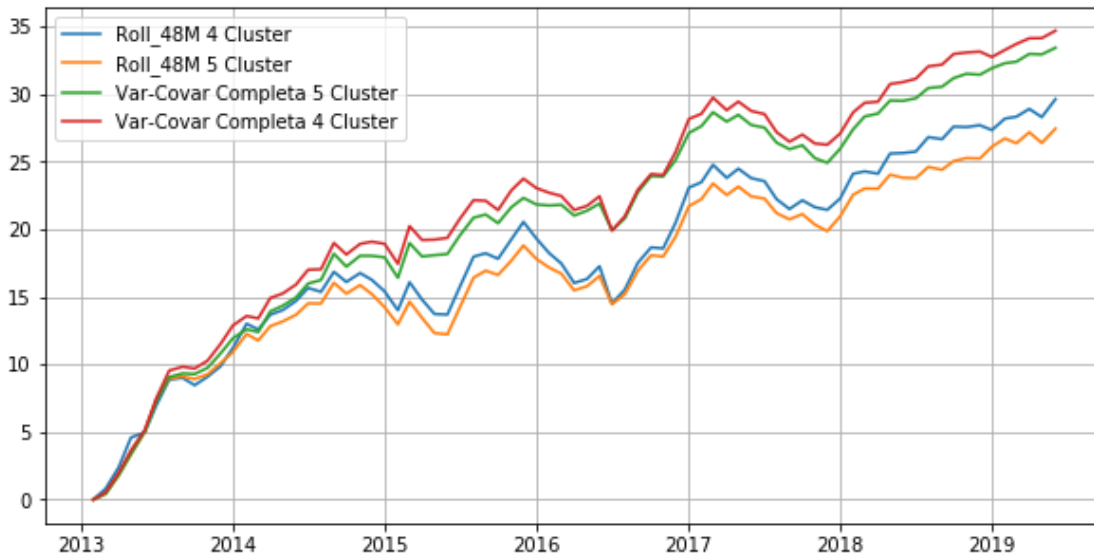


Figura 28: Rendimiento acumulado 4 fases vs 5 fases. [PB]

Dado que los rendimientos son muy parecidos, se concluye que agregar una nueva fase no agrega valor al modelo y solo lo complejiza, puesto que el resultado en ambos casos es peor. Por lo tanto, 4 fases parece ser una buena elección.

### Cambios en la Matriz de Varianza-Covarianza

Al realizar el ajuste propuesto por Husnain, Hassan y Lamarque (2016) a los portafolios con mejores rendimientos se obtienen los siguientes resultados:

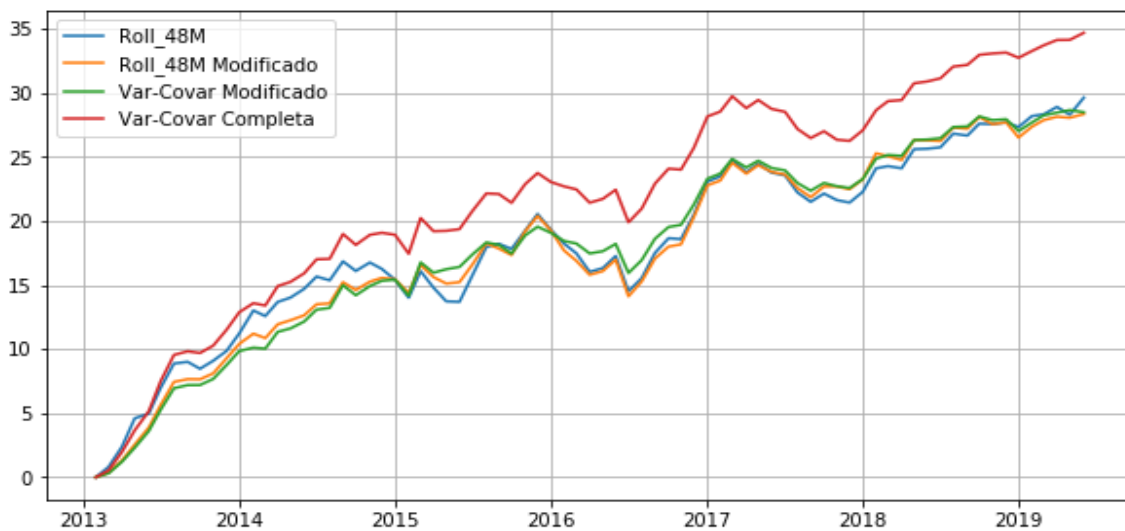


Figura 29: Rendimiento Acumulado Matriz Sin Ajustar vs Ajustada. [PB]

Como es posible evidenciar, el modelo con la matriz de varianza-covarianza ajustado tiene menor rendimiento que el sin ajustar para el caso de utilizar toda la data disponible, esto implica que el ajuste sobrestima el riesgo. Además, para el caso que se utiliza rolling de 48 meses el resultado no cambia en demasía, por lo que es posible concluir que en este caso se tiene que el aporte de la varianza de las clases de activo es mucho más significativo que el de las correlaciones en la volatilidad del portafolio<sup>62</sup>.

Por otro lado, está lo expuesto por López de Prado (2019), obteniéndose los siguientes resultados:



Figura 30: Rendimiento Acumulado Matriz Sin Ajustar vs Ajustada. [PB]

Como es posible evidenciar, esta técnica es equivalente a no colocar una volatilidad objetivo como restricción, pues las curvas están sobrepuestas. Esto se produce por el hecho de que al modificar la matriz de varianza-covarianza para la optimización del portafolio, dada las restricciones de posición, no es posible encontrar una combinación de clase de activos tal que la volatilidad sea mayor a la volatilidad objetivo. Por lo tanto, esta técnica no es útil para el problema.

<sup>62</sup> Recordar que para 2 activos se tiene que  $\sigma_p^2 = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2 * x_1 * x_2 * \sigma_{1,2}$  si los activos tienen baja covarianza la volatilidad del portafolio disminuye.

## Simulaciones

Al realizar simulaciones se busca testear el rendimiento acumulado de la estrategia, donde se analizan las dos más interesantes: Markowitz con rolling 48 meses con y sin restricción en volatilidad. Estos análisis poseen las siguientes características:

### Caso 1:

- 500 simulaciones.
- Restricciones de tamaño  $\pm 1\%$  por asset class y suma.

### Caso 2:

- 500 simulaciones.
- Restricciones de tamaño  $\pm 1\%$  por asset class y suma,  $\pm 0.25\%$  renta fija local.

## Markowitz 48 Meses Rolling Con Volatilidad Objetivo

### Caso 1:

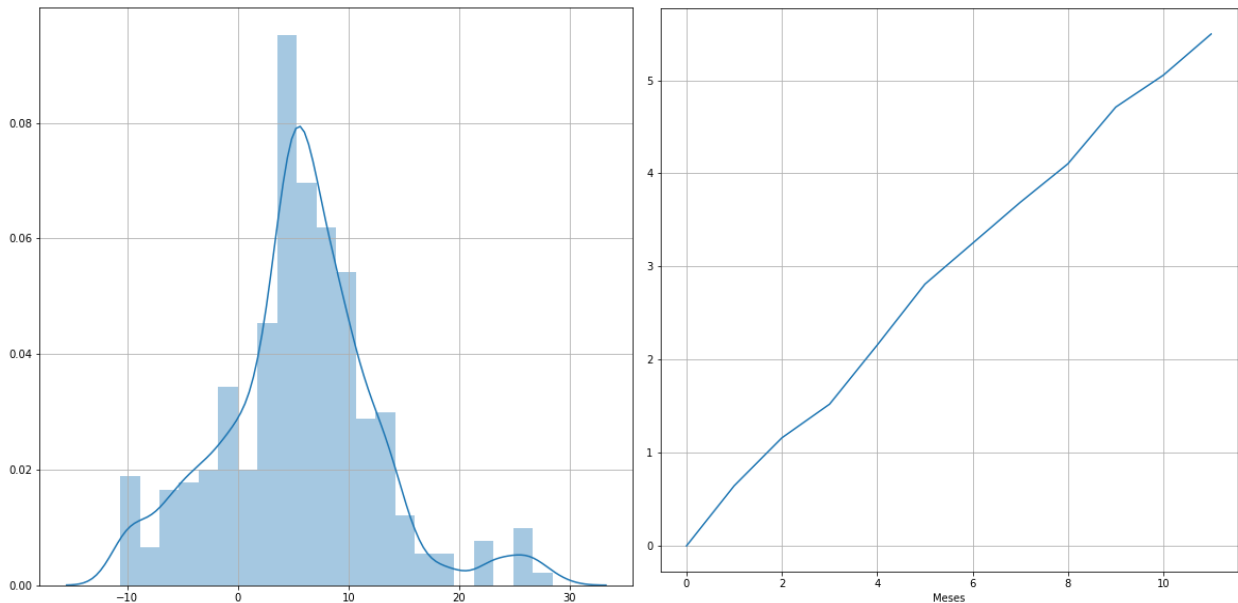


Figura 31: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 1. [PB]

## Caso 2:

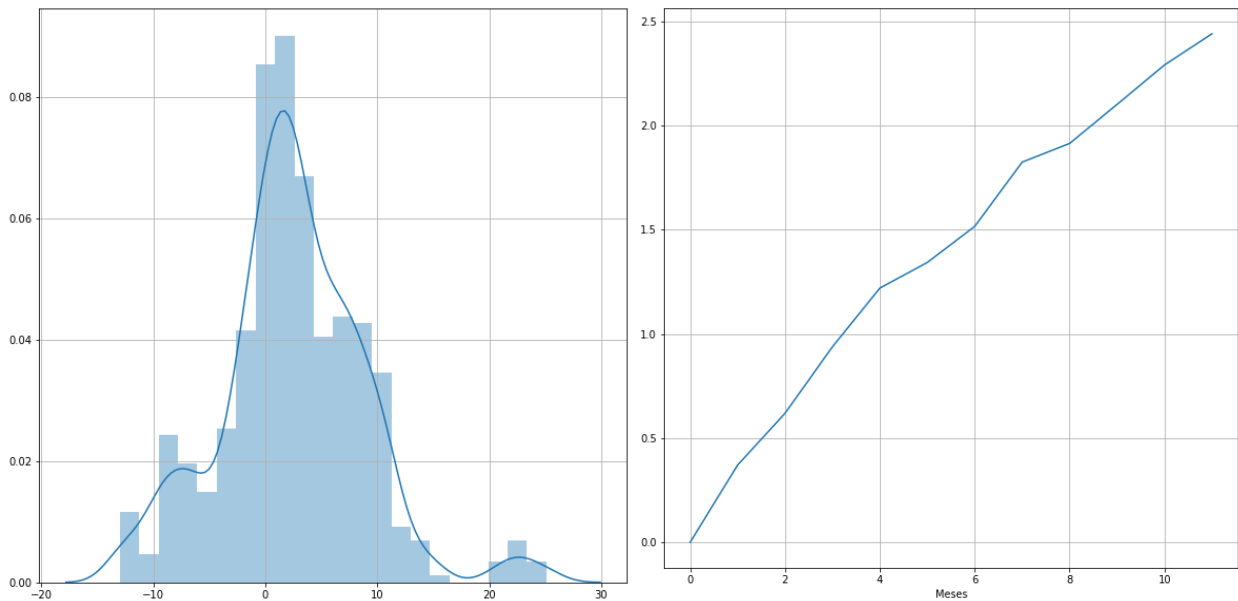


Figura 32: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 2. [PB]

Como es posible observar de las figuras, para ambos casos la distribución está centrada en un valor mayor a 0, implicando que, en promedio, la estrategia genera retorno positivo indistintamente del periodo en el que se encuentre, a su vez, es posible apreciar que es más probable que se obtengan retornos anormalmente buenos que malos, lo cual es un gran hallazgo y que se confirma por ambas curvas de retornos acumulados promedio.

Otro aspecto relevante a primera vista es la performance de la estrategia, puesto que dada la volatilidad objetivo, un performance bueno en un año es de alrededor de 7 puntos bases, el cual en el caso 1 es próximo a lograrse, sin embargo, dado los resultados obtenidos al restringir renta fija local en el caso 2, no es posible argumentar que el modelo sea capaz cumplir el objetivo del trabajo de título, puesto que la performance se ve mermado considerablemente, implicando que la esta estrategia está basada en renta fija.

También se debe considerar que los retornos anormalmente grandes se deben principalmente a inversiones en renta variable, por lo tanto, este modelo con las características señaladas si bien tiende a tomar grandes posiciones en renta fija -lo que significa, en gran medida, su performance- también toma posiciones en renta variable, implicando un espacio de mejora, como por ejemplo, restringir aún más las posiciones en los distintos activos de renta fija, para así, tener un performance más ad hoc a lo que el trabajo de título busca.

## Markowitz 48 Meses Rolling Sin Volatilidad Objetivo

### Caso 1:

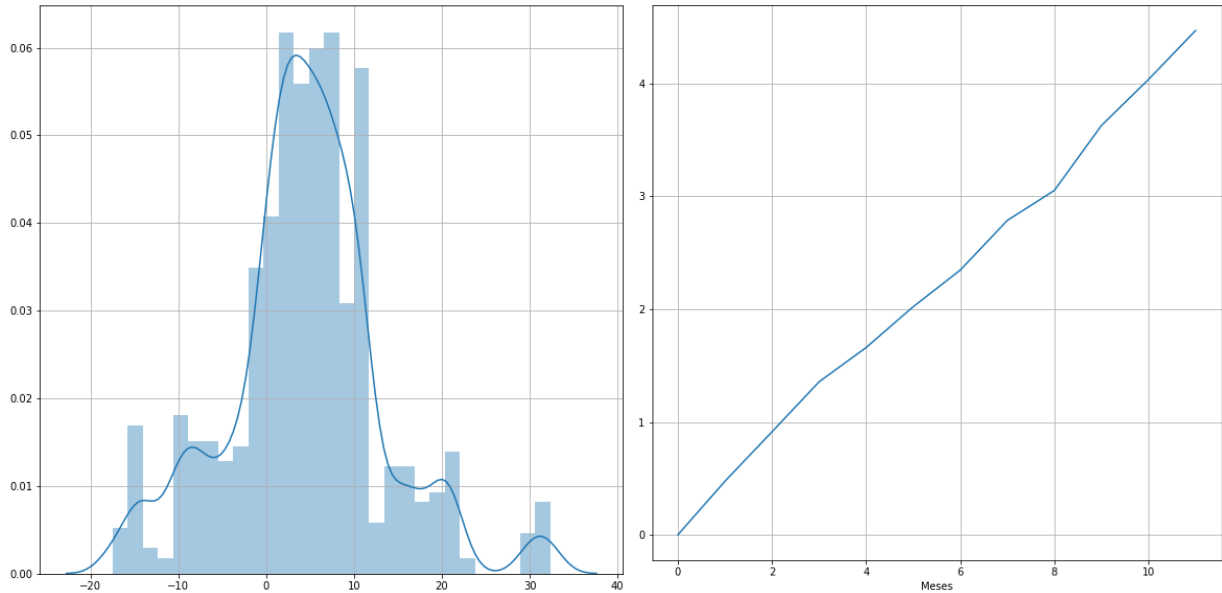


Figura 33: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 1. [PB]

### Caso 2:

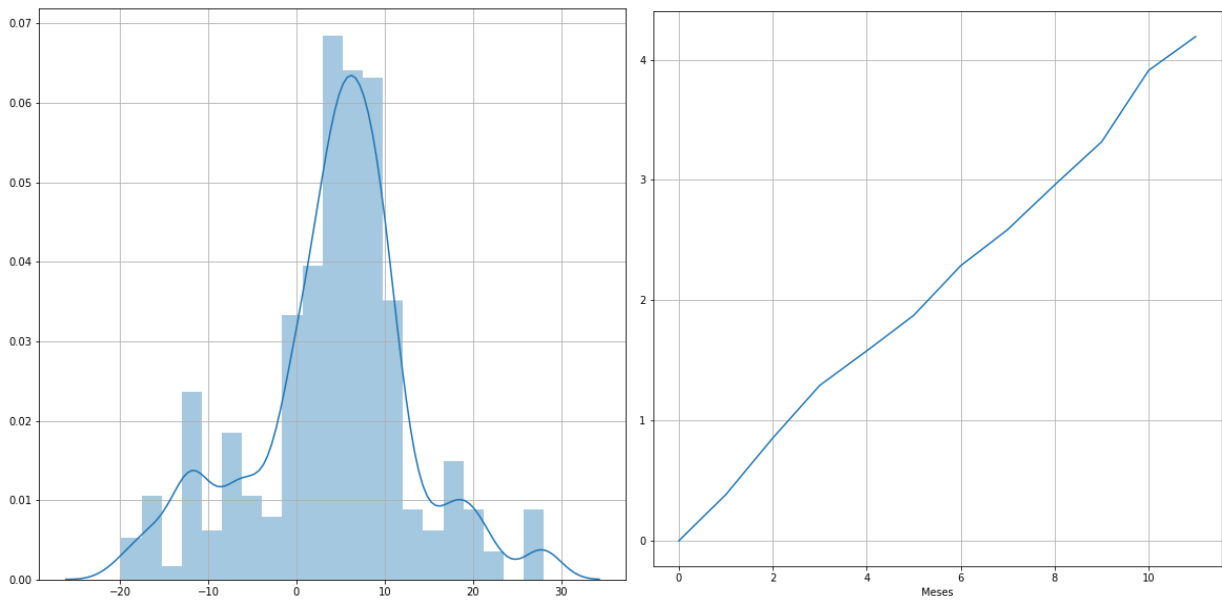


Figura 34: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 2. [PB]

Como se evidencia en las figuras, si se comparan las distribuciones del caso 1 de ambas estrategias (con y sin volatilidad objetivo), esta estrategia está centrada más cerca del 0 que la otra, resultado que viene dado porque, tal y como se observa en el análisis del modelo irrestricto, al no considerar volatilidad objetivo se toman mayores posiciones en renta variable, lo que también explica que los retornos anormales sean más grandes que en la estrategia anterior y, a su vez, que las colas de la distribución sean más parecidas. También, es posible apreciar que se espera, en general, que el retorno de la estrategia en un año sea positivo y, además, que se obtengan más retornos anormales buenos que malos.

Un aspecto interesante, es que el caso 2 de esta estrategia no es considerablemente peor que el caso 1, más aún, está centrado más hacia la derecha que este último, a su vez, la rentabilidad promedio de ambos casos parecen no ser tan distintas, por lo tanto, cuando se restringe la posición en renta fija local el modelo invierte en un activo que tuvo un retorno igual o mejor que el de renta fija local, por lo que es posible deducir que los retornos esperados por fase en otras fases están mejor definidos que para la estudiada entre 2013 y 2019.

Por otro lado, la volatilidad en el periodo estudiado dada las restricciones de tamaño son alrededor de 2.2 puntos base promedio para Markowitz, por lo que se espera que el modelo tenga una rentabilidad anual de 7.62 puntos base en un año base para considerarse bueno. Asumiendo que la volatilidad promedio es parecida en cada periodo y, considerando además que esta estrategia toma mayores posiciones en renta variable, junto con los resultados obtenidos en el caso 1 y 2, se puede inferir que una estrategia sin volatilidad objetivo es capaz de generar exceso de retorno y ser posible de implementar para la compañía, más aún, si se considera un mánager capaz de variar las restricciones de posición por clase de activo y tipo de activo (renta variable y renta fija), logrando implícitamente que estas pasen a ser views para el modelo.

#### **8.2.8. Síntesis - KM**

En base a los resultados obtenidos se tiene que la estrategia, si bien posee un performance estable y positivo en el periodo correspondiente entre 2013 y 2019, este se debe gran parte a renta fija y, en específico, renta fija local, lo cual también se refleja en las simulaciones restrictas realizadas.

A su vez, los retornos esperados por fase juegan un rol fundamental en este modelo, puesto que como se aprecia en la estrategia ranking, donde por formulación se toman posiciones más grandes en aquellos activos que se espera tengan mayores retornos, en el periodo estudiado de 2013 a 2019 no se obtiene un buen resultado, implicando que, en específico, este punto sea un aspecto para mejorar en trabajos futuros.



Dicho esto, el modelo no parece cumplir, en primera instancia, con el objetivo principal del trabajo que es generar exceso de retorno para la compañía, puesto que en el periodo estudiado la estrategia no es atractiva desde un punto de vista de mánager, dado que se basa fuertemente en renta fija local y, más aún, la performance parece depender de cierta clase de activo de manera persistente, lo cual viene determinado por los retornos esperados en la fase, puesto que como se aprecia en el apartado irrestricto, aún sin volatilidad objetivo carecía de un buen performance y tenía un perfil ligado a renta fija.

Sin embargo, dado los resultados de las simulaciones en el caso no restringido con volatilidad objetivo, es posible visualizar que el periodo en el que se realiza el análisis puede que sea perjudicial para el modelo o, en su defecto, la matriz de varianza-covarianza calculada genera una aversión al riesgo considerable y, dada los retornos esperados por fase, se fuerza a tomar mayor posición en renta fija, en específico, se encontró que la volatilidad del modelo irrestricto ronda los 2.2 puntos bases promedio, con esto en consideración y, sumando además que en este caso irrestricto en volatilidad al restringir renta fija local no hubo una baja considerable en la performance, es posible inferir que en otro periodo el modelo es capaz de generar exceso de retorno con activos más volátiles. Por lo tanto, en segunda instancia, este modelo no se debe desechar y, bajo ciertas características, puede servir como base para una mezcla de modelos.

## 8.2.9. GMM

### Caracterización de las Fases

En base a los retornos esperados por fase y, dado lo expuesto por Defend, Porteli y Cesarini (2017), donde exponen la existencia de una quinta fase, expandiendo lo mostrado en KM:

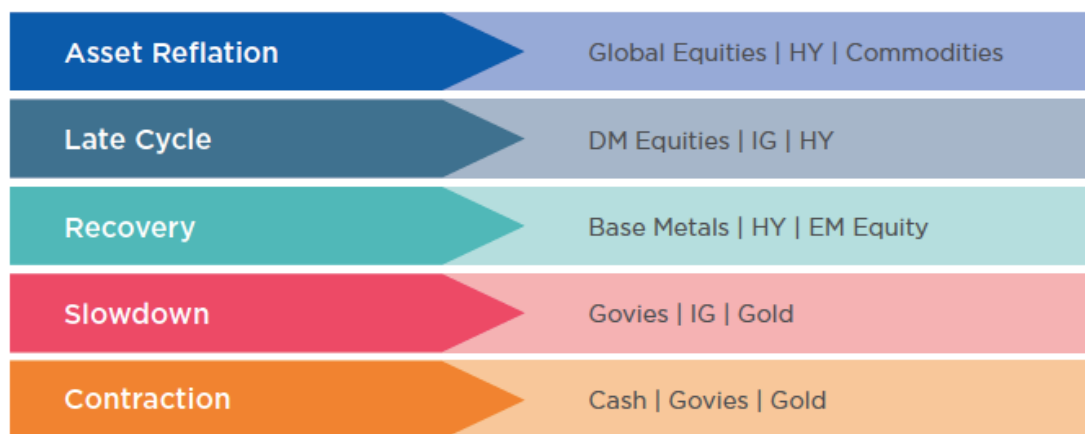


Figura 35: Características de los Ciclos. 5 fases.

Concluyéndose:

- Fase 0: Expansión.
- Fase 1: Contracción (recesión).
- Fase 2: Recuperación.
- Fase 3: Slowdown.
- Fase 4: Asset Reflation.

La reflation, según Investopedia<sup>63</sup>, hace relación a una política monetaria o fiscal que tiene como objetivo expandir la producción, estimular el gasto y frenar los efectos de deflación. Esta puede ocurrir luego de un periodo económico de incertidumbre o, también, se puede utilizar el termino en la primera fase de una recuperación económica.

<sup>63</sup> Investipedia. Reflation. Disponible en: <https://www.investopedia.com/terms/r/reflation.asp> (consulta: 8 de enero de 2020).

## Backtest

Para construir el backtest se utiliza la misma lógica empleada en KM, la que se resume a continuación:

- Periodo: 31 de enero 2013 hasta el 31 de mayo de 2019.
- Límites superiores de  $\pm 1\%$  tanto por clase de activo como por suma de tipo de activo.
- Volatilidad objetivo de 2 puntos base.

Con el siguiente periodo a backtestear:

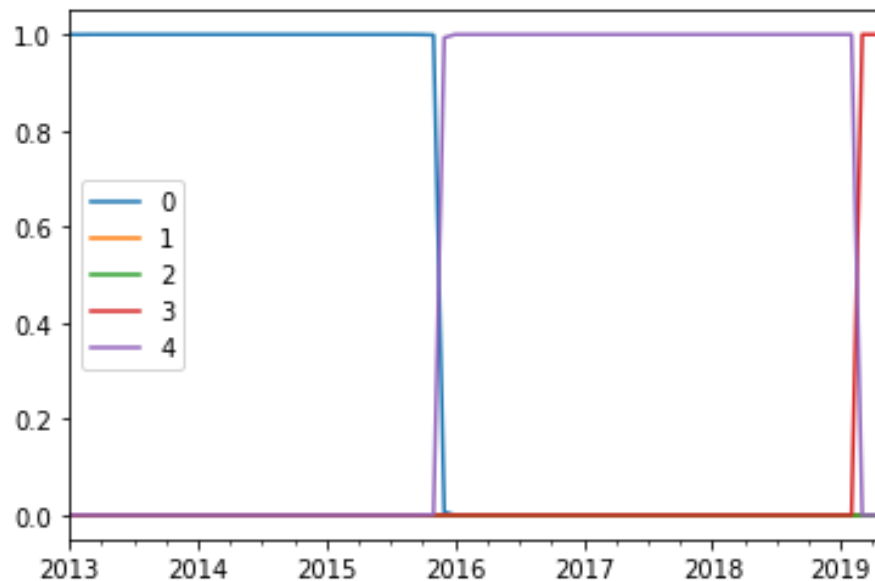


Figura 36: Fase Backtest. GMM.

Como es posible visualizar, el periodo comienza en una fase de expansión alrededor de 3 años, para luego estar otros 3 años en una fase de asset reflation, la cual, según lo expuesto, se puede interpretar como una fase de incertidumbre, por último, al final del periodo cambia de fase a slowdown. A continuación, se muestran los retornos esperados para cada una de las fases señaladas para el 31 de mayo de 2019, los cuales deben poseer el mismo signo para todo el intervalo a estudiar:

	Expansión	Slowdown	Asset Reflation
<b>.DMEQLCEW Index</b>	5.37%	-5.05%	2.02%
<b>.EMEQCULC Index</b>	0.57%	3.57%	3.81%
<b>.CHFICUAA Index</b>	3.86%	4.50%	3.34%
<b>LUATTRUU Index</b>	1.19%	3.94%	1.51%
<b>.HBFIWOCO Index</b>	2.29%	10.29%	3.42%

Tabla 14: Retornos Esperados GMM.

Markowitz:

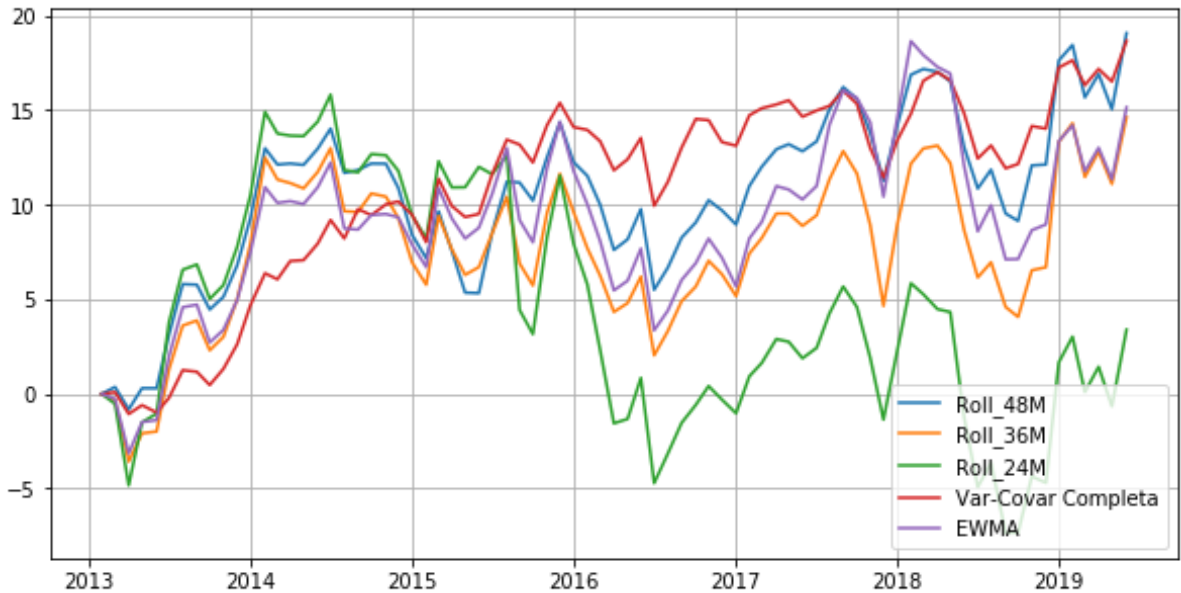


Figura 37: Rentabilidad Acumulada Markowitz. [PB]

Como es posible evidenciar, el perfil de la estrategia es considerablemente más volátil que en el caso de KM, más aún, entre menos data se considera para calcular la matriz de varianza-covarianza más se evidencia una curva volátil, lo que implica que se toman mayores posiciones en activos como renta variable. Por otro lado, si bien la performance es menor que en el caso de KM con las mismas restricciones, la forma de la curva es más atrayente que la de este último caso, principalmente por el hecho de que una curva con perfil de renta fija, como lo encontrado KM, no es atractiva para un manager de inversión, por el hecho de que al ser un activo con bajo retorno real no es posible cumplir con el principal objetivo de la compañía que es generar diferenciación en base a mayores retornos. A continuación, se visualiza las dos estrategias con mejor performance:

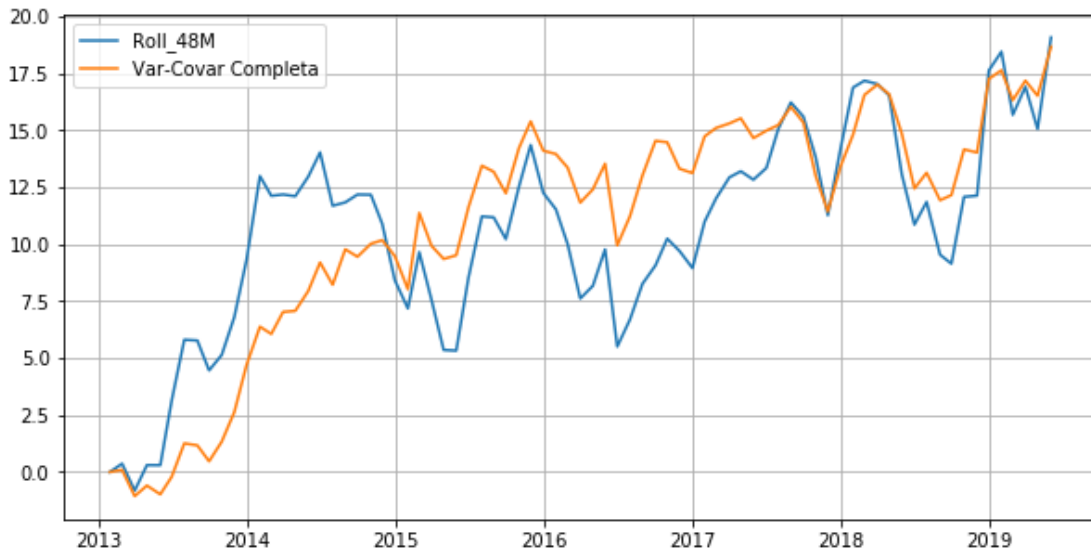


Figura 38: Mejor Performance Markowitz. [PB]

La figura es sumamente interesante, especialmente la curva roll de 48 meses. Por un lado, se encuentra la curva relacionada a la matriz de varianza-covarianza utilizando la data histórica completa que, si bien presenta un perfil volátil, posee menores variaciones y un retorno similar al de Rolling 48 meses que, por otro lado, presenta una forma mucho más volátil. De esta última curva el hecho visualmente más relevante es que las ganancias se generan en los periodos comprendidos entre 2013 y mitad del 2014, y mediados de 2018 hasta el 31 de mayo de 2019, pues el periodo comprendido entre 2014 y mediados de 2018 se puede considerar como “flat”, lo que se traduce como que las variaciones que va teniendo el portafolio en ese periodo se terminan anulando (tanto en ganancia como en pérdida). Lo anterior se puede verificar comparando los portafolios de las estrategias:

Index	DMEQLCEW Index	EMEQCULC Index	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	-12.7891	-9.42298	70.9961	-100	51.216
std	36.727	31.9332	24.7738	1.16526e-13	39.8946
min	-100	-45.9623	25.4448	-100	-3.11499
25%	-47.158	-36.3137	49.1321	-100	7.716
50%	-15.9334	-27.9911	70.5436	-100	58.2801
75%	27.1972	15.5977	100	-100	97.184
max	30.7405	83.5861	100	-100	100

Figura 39: Caracterización Portafolio Markowitz Var-Covar Completa. [PB]

Index	DMEQLCEW Index	EMEQCULC Index	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	-13.0579	-10.4866	57.6865	-100	65.8579
std	47.4838	61.4795	38.103	1.35318e-13	42.9349
min	-97.2554	-100	-12.998	-100	-9.12651
25%	-55.3311	-73.6157	29.9366	-100	18.0942
50%	-4.11404	-44.4983	61.5941	-100	98.3068
75%	34.8539	35.2357	100	-100	100
max	52.2799	98.0383	100	-100	100

Figura 40: Caracterización Portafolio Markowitz Rolling 48 M. [PB]

Tal y como se evidencia, las posiciones en renta variable para la estrategia Rolling son en promedio similares a la var-covar completa, sin embargo, la desviación estándar es considerablemente distinta, lo que implica que toma posiciones más extremas, lo que explica el perfil volátil visualizado, por otro lado, sistemáticamente se evidencia que el algoritmo toma la máxima posición corta en renta fija investment grade, lo que se subentiende como que utiliza esta clase de activo para financiarse y, por lo tanto, podría ser reemplazado por el manager a cash extranjero.

Por otro lado, se tiene que renta fija local es un mercado más pequeño y, por consiguiente, es menos líquido. Por lo que se realiza el análisis restringiendo su posición que, según juicio de experto, 0.25% parece ser un buen valor, obteniéndose los siguientes resultados:

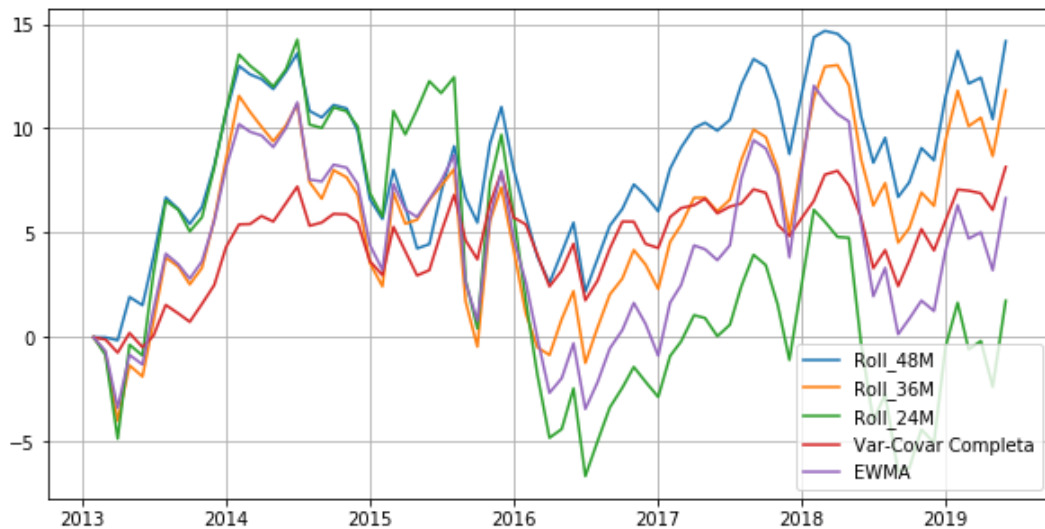


Figura 41: Rentabilidad Acumulada RFL Restringido. [PB]

Si bien la performance disminuye, esta disminución no es en la misma magnitud que en KM, lo que también confirma el hecho de las posiciones en renta fija local para esta estrategia es, per se, menor. Por otro lado, se tiene que Rolling 48 meses vuelve a destacar y que var-covar completa disminuye su performance, resultado esperado dado que es la estrategia que mayor posición toma en renta fija (considerando su aversión al riesgo al tomar la data histórica para calcular la matriz de varianza-covarianza).

## Ranking

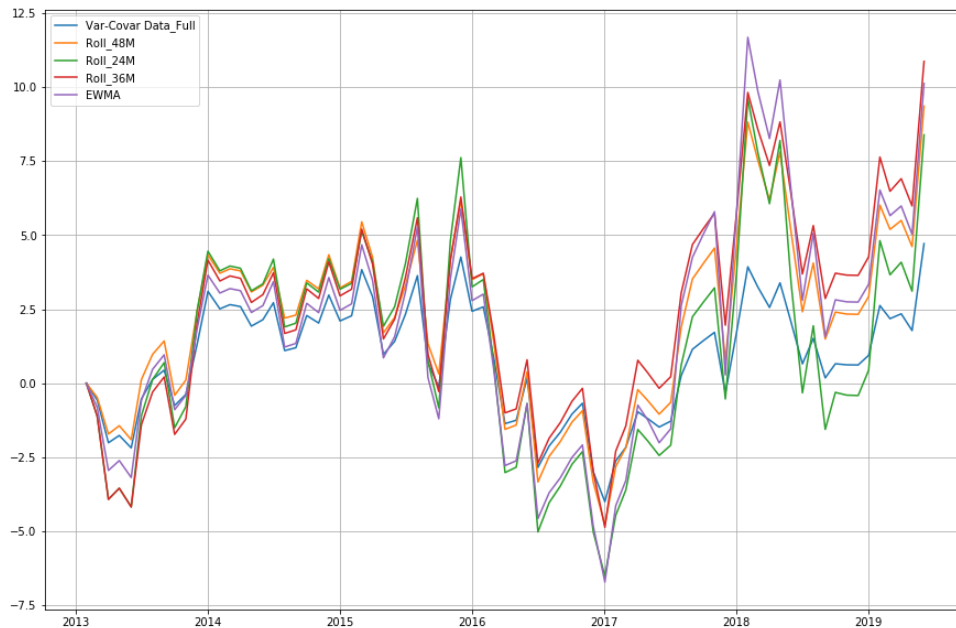


Figura 42: Rentabilidad Acumulada Estrategia Ranking. [PB]

Al igual que en KM, la performance de la estrategia disminuye considerablemente para cada una de las combinaciones con respecto a Markowitz, por otro lado, como se ha evidenciado en cada uno de los resultados restringidos en volatilidad, se aprecia una disminución considerable en el año 2016, además, se aprecian curvas volátiles considerables, implicando posiciones más grandes en renta variable. A su vez, se tiene que el bajo retorno acumulado de este modelo puede venir dado por su estructura, donde como se identifica en KM, se puede deber a dos razones:

- La regresión lineal no es buena función para caracterizar los retornos esperados por fase.
- Dado que la formulación obliga a tomar posiciones grandes donde se espera tener mejores resultados, es posible que el tamaño de las posiciones grandes sea en sí pequeños.

Para evidenciar lo anterior se analizan las dos estrategias con mayor performance para ranking, Rolling 36 meses y EWMA.

Index	DMEQLCEW Inde:	EMEQCULC Inde>	.CHFICUAA Index	LUATTRUU Index	.HBFIWOCO Index
count	77	77	77	77	77
mean	5.16596	-0.532827	11.7702	-48.4826	32.0793
std	46.5765	57.2363	16.4453	27.4061	30.7947
min	-61.6503	-99.7988	-26.7789	-99.9	2.60381e-150
25%	-35.1952	-52.1054	5.46329e-179	-68.4296	4.92758e-50
50%	-4.67914e-1...	-31.3226	9.52605e-26	-48.7996	33.8589
75%	52.23	67.104	26.618	-26.4028	49.8429
max	64.9885	99.7219	34.4771	-8.94181e-2...	99.9

Figura 43: Caracterización Portafolio Estrategia Rolling 36M. [PB]

Index	DMEQLCEW Inde:	EMEQCULC Inde>	.CHFICUAA Index	LUATTRUU Index	.HBFIWOCO Index
count	77	77	77	77	77
mean	4.10972	4.46775	11.9098	-53.5766	33.0893
std	49.4213	63.6047	18.2074	32.1382	29.2458
min	-60.2239	-69.2689	-38.8205	-100	2.35053e-150
25%	-49.9188	-53.3118	5.6114e-179	-76.6404	4.87652e-50
50%	-3.98528e-1...	-33.6268	9.19405e-26	-34.6345	36.2178
75%	53.8037	70.3359	28.4033	-27.6674	50
max	71.8008	100	35.9004	-7.74684e-2...	100

Figura 44: Caracterización Portafolio Estrategia EWMA. [PB]

Como es posible apreciar, para ambos casos se tiene que las posiciones en renta variable son mayores (en términos de tamaño) con respecto a renta fija, lo que explica la performance volátil evidenciada. Por otro lado, las posiciones en sí no se pueden catalogar como pequeñas, puesto que, dada la restricción de volatilidad, toma posiciones grandes en renta variable, por lo tanto, el bajo performance se debe, principalmente, a una mala caracterización de los retornos esperados por fase, en específico, en el periodo comprendido entre 2016 a 2017 y 2018 a 2019, que implicaron pérdidas cuantiosas de 15 y 12 puntos bases respectivamente, pues en el resto del periodo se evidencia ganancias.



Es interesante destacar que sistemáticamente se tiene una baja en la performance en el periodo comprendido entre 2016 y 2017 independiente de la combinación, similar a lo que pasaba con el modelo KM y, además, se suma el periodo comprendido entre 2018 y 2019, hechos que implican una merma considerable en la rentabilidad de la estrategia, por otra parte, no resulta relevante analizar el hecho de cerrar las posiciones con caja, puesto que es prácticamente imposible que este efecto amplifique la rentabilidad a un nivel considerado bueno de 44 puntos bases. Por la misma razón, tampoco se analizan combinaciones del tipo roll-ewma.

Otra característica diferenciadora de este modelo viene dada por la forma de la curva, en la que es posible apreciar un perfil volátil considerablemente mayor al que se vio en KM, sea cualquiera la combinación. La principal razón de esto viene dada, principalmente, por las características de GMM, en primer lugar, es posible evidenciar que las probabilidades de pertenencia a las fases son cercanas a 1, implicando modificaciones en la matriz de varianza-covarianza y los retornos esperados por fase que, tal y como se aprecia, influyen en tomar posiciones mayores en activos más riesgosos.

Por último, que las probabilidades sean cercanas a 1 en todo periodo de una fase implica que los portafolios van a ser similares o iguales durante ese tiempo, por lo que solamente ocurre un cambio cuando se pasa a otra fase, implicando que las posiciones a lo largo del periodo de una fase solamente se van amplificando en las clases de activo. Esta característica es negativa si se considerara la estrategia como única para generar allocation, sin embargo, si se busca hacer mezcla de modelos, el allocation de GMM se puede entender como una directriz de posiciones en base a la fase del ciclo económico en que se esté, que a su vez, se puede complementar con otro modelo capaz de capturar otras características.

## 8.2.10. Análisis de Sensibilidad - GMM

### Modificación de la Volatilidad Objetivo

Tal y como para el modelo KM, se analiza el modelo irrestricto en volatilidad, es decir, sin considerar volatilidad objetivo. Considerando los límites establecidos en un inicio por clase de activo y suma de tipo (renta variable y renta fija)  $\pm 1\%$ .

Markowitz:

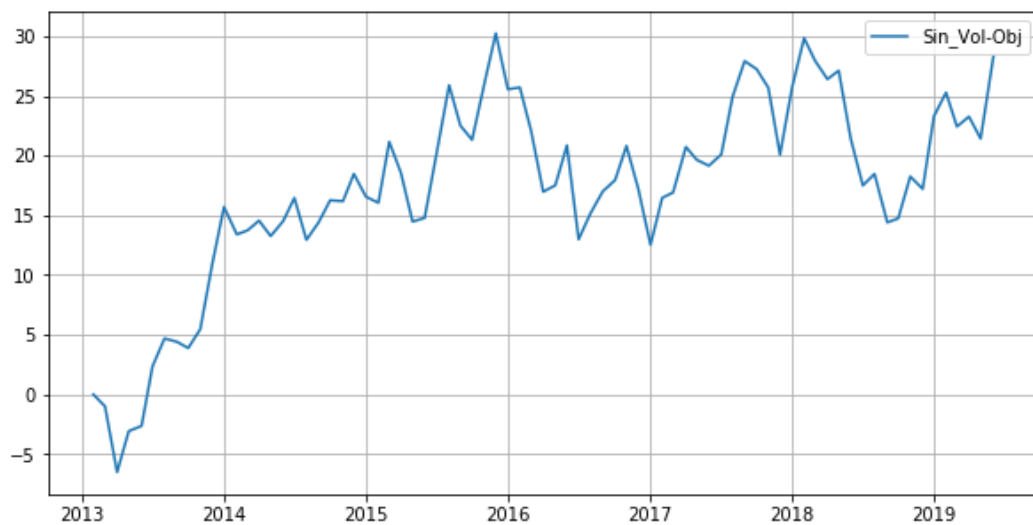


Figura 45: Rentabilidad Acumulada Estrategias Markowitz Sin Vol-Obj. [PB]

En primer lugar, se aprecia un performance más volátil que en el caso restringido, lo cual es esperable dado que no se restringe las posiciones a una volatilidad objetivo, además, se evidencia un periodo “flat” entre 2014 y mediados de 2018, por lo que en este caso irrestricto las ganancias se generan entre los periodos 2013-2014 y mediados 2018-2019. Dicho lo anterior, es interesante analizar las volatilidades considerando la matriz rolling 24 meses (menos aversión al riesgo) y var-covar completa (más aversión al riesgo):

<b>Volatilidad [BP]</b>	<b>Roll 24 M</b>	<b>Var-Covar Completa</b>	<b>Promedio</b>
<b>Promedio</b>	3.27	5.14	4.205
<b>Mediana</b>	3	5.02	4.01
<b>Minimo</b>	1.43	3.78	2.605
<b>Maximo</b>	4.97	8.22	6.595

Tabla 14: Volatilidad Portafolio Sin Vol-Obj. [PB]

Los resultados obtenidos vienen a confirmar lo expuesto y, mejor aún, generan un hallazgo. En primer lugar, para ambos perfiles (alta y baja aversión al riesgo) se tiene que la volatilidad es mayor al objetivo propuesto por el mánager (2 puntos base), implicando que al eliminarlo se tomen mayores posiciones en, según se observa, renta variable. Con esto en consideración, se halla que si bien los límites poseen implícitamente una volatilidad, son los retornos esperados por fase los que influyen más en esta, pues para KM y GMM se tienen mismos límites, pero volatilidades muy distintas.

Por otro lado, es interesante apreciar si en este caso se tiene que mayor riesgo implica mayor retorno, por lo tanto, se compara la performance con respecto a la estrategia rolling 48 meses con 2 puntos bases de volatilidad objetivo:

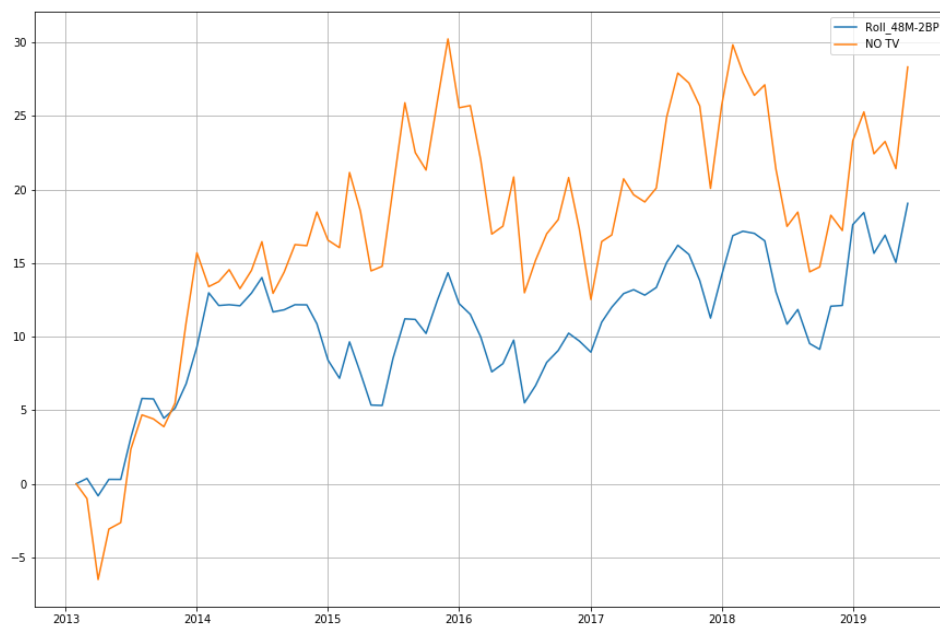


Figura 46: Rentabilidad Acumulada Rolling 48M 2 BP Vol-Obj v/s Sin Vol-Obj. [PB]

En este caso, a diferencia de KM, mayor riesgo sí implica mayor retorno, aún así no se puede inferir que la caracterización de los retornos por fase sea mejor en GMM o, en su defecto, el modelo GMM es mejor que KM (considerando los valores de las matrices de varianza-covarianza y las probabilidades), si no que se necesita un análisis más exhaustivo de otras métricas de medición, el cuál no agrega valor realizar dado los bajos resultados obtenidos para ambos modelos hasta el momento.

### Ranking

No agrega valor a la discusión analizar esta estrategia irrestricta, puesto que, dada las características de su formulación, se espera lo ya visto en su símil del modelo KM, donde la curva de rentabilidad se amplifica ya que no existe volatilidad objetivo.

## Quitar Una Fase

El quitar una fase hace relación a lo encontrado en KM y a lo que comúnmente se conoce como ciclo económico, el cual está compuesto por las cuatro fases ya mencionadas. Al comparar los resultados de las estrategias con mejor performance y perfil (roll 48 meses y sin restricción de volatilidad) se tiene que:

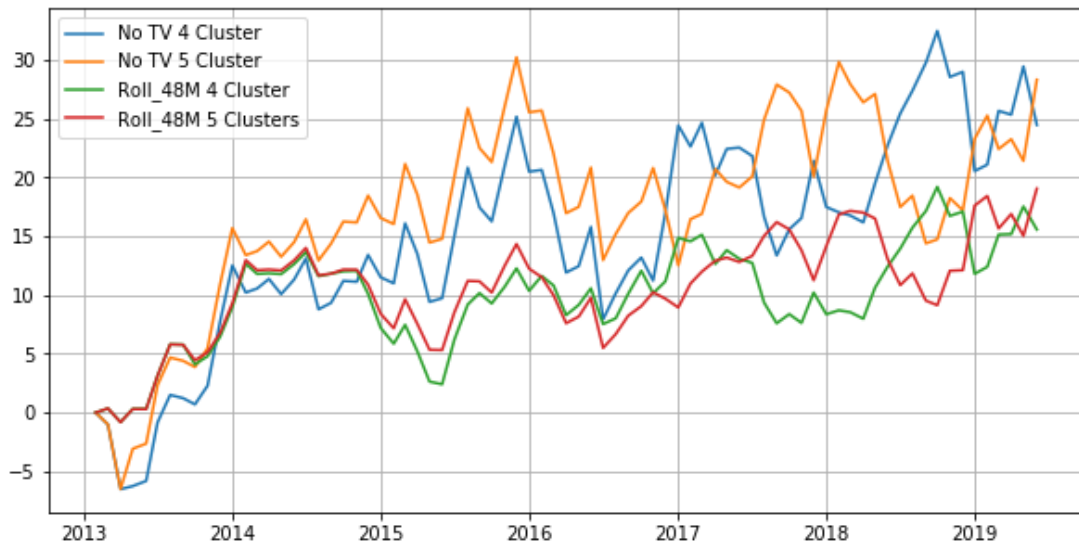


Figura 47: Rentabilidad Acumulada 5 Clusters v/s 4 Cluster. [PB]

De la figura se desprende que, en primer lugar, la performance general en el periodo estudiado es relativamente parecido, sin embargo, lo que llama más la atención es que cuando se consideran 5 fases el periodo “flat” es visualmente más evidente (2014-2018). Con esto en consideración, es más atractivo implementar el modelo GMM con 5 fases si es que se hacen mezclas de modelos.

## Modificar Matriz de Varianza-Covarianza

Dado los resultados obtenidos, donde el perfil de las distintas estrategias muestra características más volátiles que las relacionadas al modelo KM, junto con los hallazgos encontrados al analizar el modelo irrestricto, donde se determina que las restricciones de tamaño son capaces de definir, en promedio, 4 puntos bases de volatilidad, sumado además lo encontrado en el modelo KM, donde se evidencia que la técnica propuesta por Husnain, Hassan y Lamarque (2016) sobreestima el riesgo, implicando menor performance y un perfil asociado más a renta fija que, tal como se ha dicho, no es de interés, y por otro lado que lo expuesto por López de Prado (2019) genera un cambio tal en la matriz de varianza-covarianza que dada las restricciones de tamaño existentes es equivalente a no considerar una volatilidad objetivo, se tiene que no es relevante realizar este análisis para el modelo GMM puesto que en ninguna circunstancia se alcanza el objetivo propuesta en el trabajo de título.

## Simulaciones

Al realizar simulaciones se busca testear el rendimiento acumulado de la estrategia, donde se analizan las dos más relevantes: Markowitz con rolling 48 meses con y sin restricción en volatilidad. Estos análisis tienen las siguientes características:

### Caso 1:

- 500 simulaciones.
- Restricciones de tamaño  $\pm 1\%$  por clase de activo y suma (renta fija y renta variable).

### Caso 2:

- 500 simulaciones.
- Restricciones de tamaño  $\pm 1\%$  por clase de activo y suma (renta fija y renta variable),  $\pm 0.25\%$  renta fija local.

## Markowitz 48 Meses Rolling Con Volatilidad Objetivo

### Caso 1:

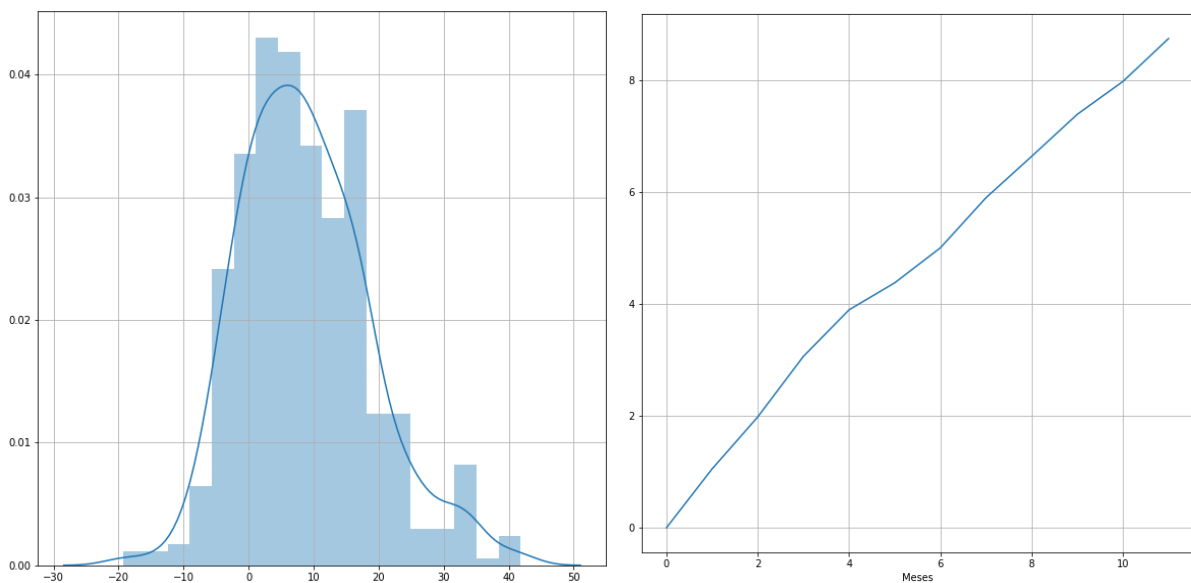


Figura 48: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 1. [PB]

## Caso 2:

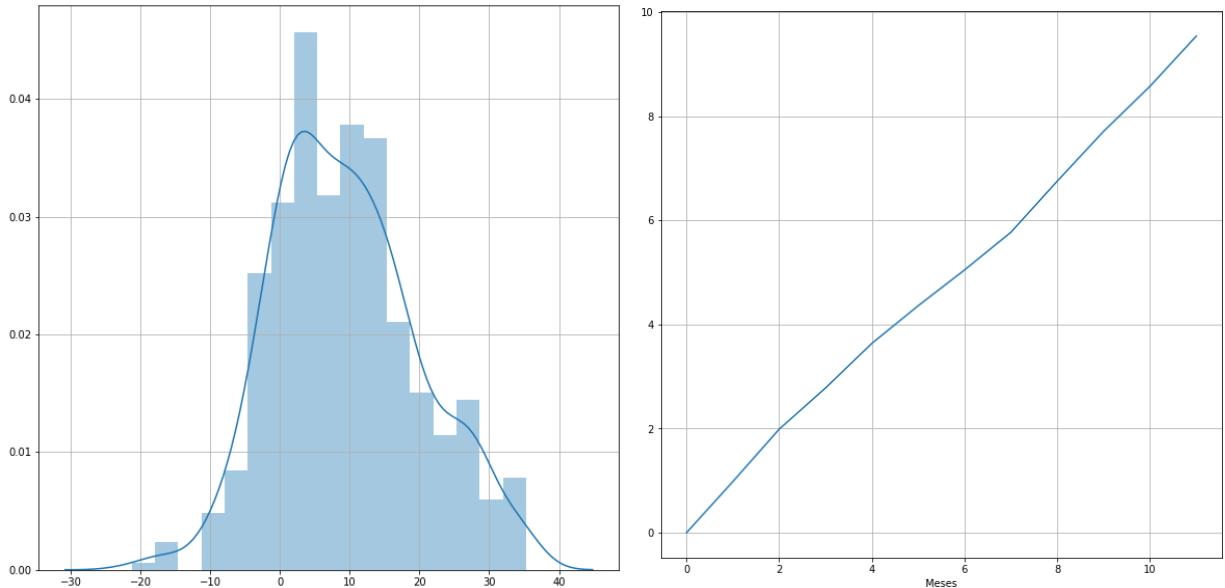


Figura 49: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 2. [PB]

Como es posible apreciar en las figuras, los hallazgos son mejores que en el modelo KM, pues la distribución se encuentra centrada más a la derecha del 0 en ambos casos y, lo más llamativo de todo, es el hecho de que rendimientos acumulados anormalmente positivos son considerablemente más probables que los anormalmente malos, condición que en KM no ocurre de la misma manera. Es importante recordar que estos retornos anormalmente mayores vienen dados, principalmente, por renta variable, dada su mayor volatilidad, por consiguiente y, tal como se aprecia, la forma de estas distribuciones comparadas con sus símiles del modelo KM es más achatada y con valores más extremos.

Por otro lado, un aspecto relevante es el hecho que, al tener una volatilidad objetivo de 2 puntos bases, se espera que en un año se obtengan alrededor de 7 puntos base de retorno (para considerar el modelo bueno), cantidad que se ve ampliamente superada y, más llamativo aún, en el caso de restringir renta fija local se tiene que la performance promedio es aún mejor que en el caso con renta fija local con límite de 1%, por lo que cuando toma posiciones en distintos activos al acotar renta fija local, se tiene que estos rentan de buena manera.

Con estos resultados y, dada las características ya mencionadas en el análisis del modelo, se tiene que el modelo GMM restringido con volatilidad objetivo puede ser una buena directriz de movimiento que, acompañado de un buen complemento, puede generar exceso de retorno persistente, junto con ser, además, posible de implementar para la AFP.

Markowitz 48 Meses Rolling Sin Volatilidad Objetivo

Caso 1:

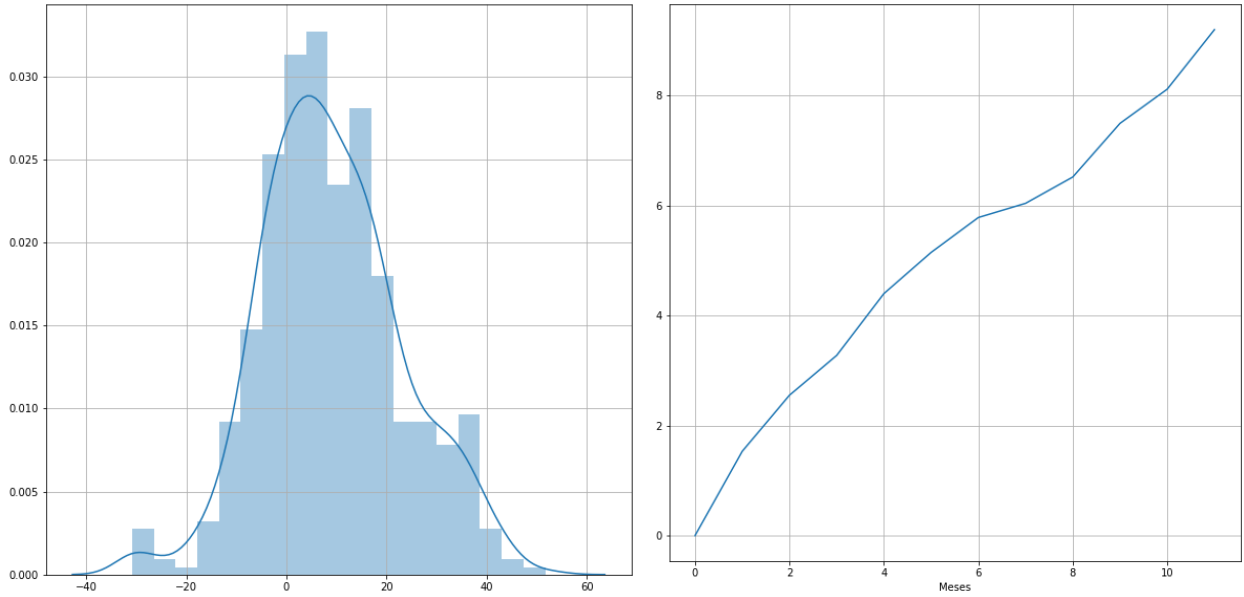


Figura 50: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 1. [PB]

Caso 2:

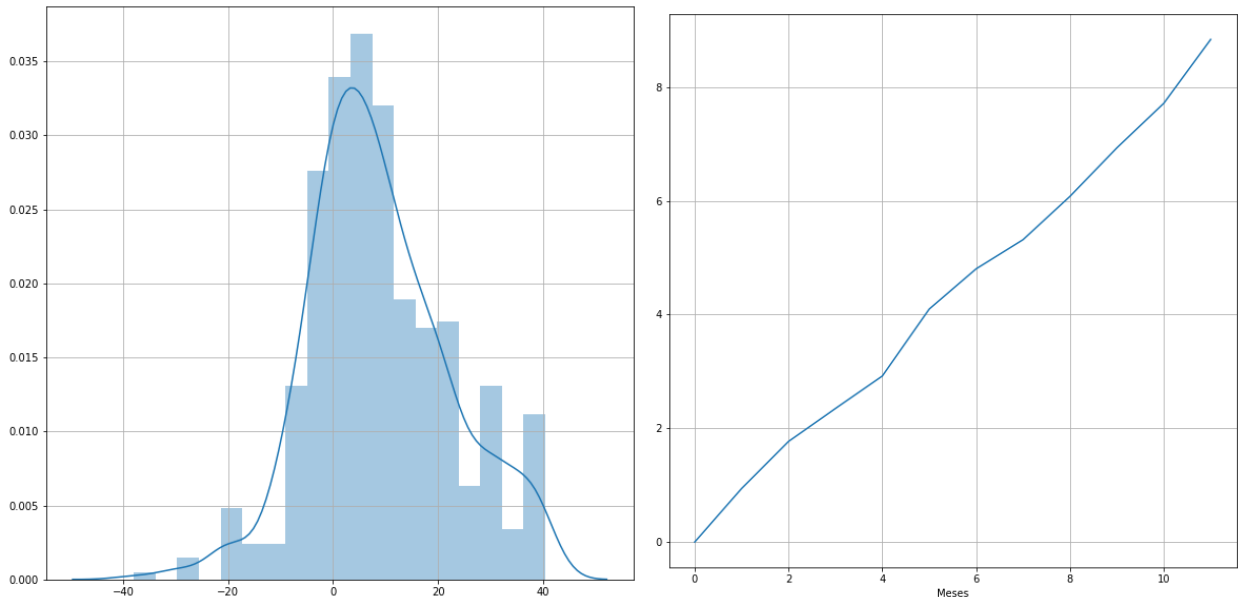


Figura 51: Distribución Retornos acumulados y Retorno Acumulado Promedio Caso 2. [PB]

Como se puede apreciar en las figuras, existen características similares a las encontradas en la estrategia anterior del modelo GMM, diferenciándose por el hecho de que la cola de la distribución asociada a rendimientos anormalmente buenos es considerablemente más grande que la asociada a rendimientos anormalmente malos (comparado también con las estrategias equivalentes al modelo KM), sin embargo, también se aprecia que las distribuciones, para ambos casos, se encuentran centradas más próximas al 0 que en las estrategias equivalente en el modelo KM, lo que se debe, en primer lugar, por el hecho de que como se observa a través del análisis, este modelo tiende a tomar posiciones con activos más volátiles per se y, al no acotar la volatilidad, se espera que tenga retornos menos persistentes y en rangos mayores, característica que se aprecia al ver la altura de la distribución, donde se tiene que es más aplanada que en su símil del caso KM, pero con valores más extremos.

Según el análisis realizado, la volatilidad del caso irrestricto ronda los 4.2 puntos base promedio mensual, por lo que para que la estrategia sea considerada buena se espera que acumule en un año alrededor de 14 puntos base, por lo tanto, para ambos casos se tiene rendimientos acumulados promedio relativamente moderados y, también, levemente más bajos que en la simulación restringida de GMM, lo que se explica debido a que el no tener una volatilidad objetivo que restrinja las posiciones conlleva a eventos mucho más extremos, que implican que, en promedio, se tengan menores retornos, pero eventos anormalmente mejores.

Por lo tanto y, dado el análisis realizado, el modelo parece generar retorno diferenciador bajo las condiciones impuestos en esta estrategia y, a su vez, parece ser - al igual que la estrategia anterior- una buena directriz de movimientos que, complementada de buena manera, puede generar exceso de retorno persistente para la compañía.

#### **8.2.11. Síntesis - GMM**

En base a los resultados obtenidos, si bien en el periodo estudiado se tiene un performance menor que con el modelo KM, el perfil de la estrategia es mucho más llamativo para un mánager de inversión, principalmente por el hecho de poder generar retorno diferenciador a través de activos de renta variable que, tal y como se observa, viene determinado por los retornos esperados por fase y la estructura del modelo. Por otro lado, si se considera que en bastantes estrategias hay un largo periodo en el que la rentabilidad del modelo se encuentra “flat”, se tiene una oportunidad de mejora la cual puede ser, por ejemplo, utilizar otra técnica para determinar los retornos esperados por fase que, como en KM, son un elemento fundamental para el buen resultado del modelo.



A su vez, el perfil de este modelo es considerablemente más volátil que KM, lo que se explica con las mayores posiciones encontradas en renta variable al analizar los portafolios, sin embargo, la performance de la estrategia en el periodo estudiado no parece ser tal para cumplir con el objetivo general de este trabajo, puesto que en sí y, a pesar de tomar mayor posición en renta variable, no es capaz de generar diferenciación teniendo un performance poco sobresaliente.

Sin embargo, según el análisis de las simulaciones realizadas, es posible razonar que el modelo sí es capaz de generar exceso de retorno considerable en otros periodos, además, tomando en cuenta las características de este algoritmo, donde las probabilidades en cada fase son cercanas a 1, implicando que en una fase se toma constantemente una posición similar, GMM puede tomarse como una buena directriz de posiciones para generar exceso de retorno, más aún, si se considera que tanto KM como GMM son modelos que miran al largo plazo, puesto que, por ejemplo, consideran el retorno esperado a 6 meses para tomar posición. Por lo tanto, parece un buen complemento para algún otro modelo.

### **8.3. Modelo Momentum**

Para desarrollar este modelo se utiliza lo publicado por Lim, Roberts y Zohren (2019) en “Enhancing Time Series Momentum Strategies”.

#### **8.3.1. Data**

Al tratarse de un modelo que estudia la tendencia de los activos financieros, en este caso, los activos en los que se genera allocation, se utiliza como data los índices expuestos en el apartado 8.1. A su vez, para conservar la consistencia entre los modelos desarrollados, la data comienza desde el 1 de enero de 2000 hasta el 31 de mayo de 2019.

Si bien se conserva el hecho de generar allocation mensual el tercer lunes de cada mes y, de esta manera, permitir que los modelos sean comparables, para el desarrollo es necesario utilizar data diaria.

#### **8.3.2. Tratamiento de Datos**

Al igual que para el modelo de ciclos económicos, se realiza un análisis exploratorio de la data en busca de valores faltantes y outliers que pueden afectar el resultado del modelo. No se realiza un análisis de correlaciones puesto que cada uno de los índices representa a un tipo de activo en el que se genera allocation.

## Datos Faltantes

Index	Datos Faltantes
LVACLI Index	289
LD20TRUU Index	0
.DMEQLCEW Index	1
.EMEQCULC Index	2
.HBFIWOCO Index	554
.CHFICUAA Index	289
LUATTRUU Index	0

Tabla 15: Datos Faltantes.

Como es posible apreciar en la tabla, existen datos faltantes para casi todos los índices, sin embargo y, tal como se expone en el modelo de ciclos económicos, estos datos corresponden al comienzo de la serie de tiempo donde no existe información, a su vez, no es necesario eliminar ni reemplazar los datos puesto que, dada las características del modelo, no es un inconveniente relevante.

## Outliers

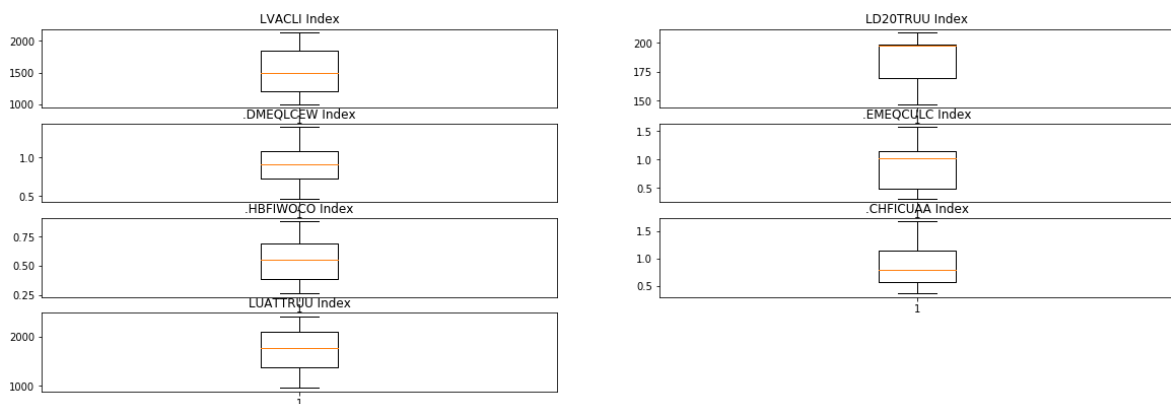


Figura 52: Outliers.

De la figura es posible evidenciar que no existen outliers en la data a utilizar, por lo tanto, se espera que no haya distorsión en los resultados pues no existen valores fuera de un rango esperable que incidan el modelo.

Por lo tanto, se decide no eliminar ninguna entrada de la base de datos. Por consiguiente, la data utilizada es la siguiente, la cual posee 5.326 entradas:

Tipo Data	Data
Índices	.DMEQLCEW Index
	.EMEQCULC Index
	.CHFICUAA Index
	LUATTRUU Index
	.HBFIWOCO Index

Tabla 16: Datos Faltantes.

### 8.3.3. Señal – Tendencia

Para generar la señal que luego se utiliza para construir el portafolio final se consideran los índices expuestos en el punto anterior. Con esto en consideración, el indicador de tendencia para el activo “i” en el tiempo “t” que propone la publicación y que es utilizado en este trabajo de título es la siguiente:

$$Y_t^i = \frac{q_t^i}{std(p_{t-252;t}^i)}$$

Donde:

$$q_t^i = \frac{MACD(i, t, S, L)}{std(p_{t-63;t}^i)}$$

A su vez,  $std(p_{t-63;t}^i)$  corresponde a la desviación estándar del precio del activo “i” considerando los últimos 63 días. También se establecen los valores de S y L como las siguientes tuplas: (8,24), (16,48) y (32,96). Por otro lado, tal como exponen Asness, Moskowitz y Pederson (2013), la señal de momentum óptima se puede generar considerando distintos periodos de tiempo que van de 3 a 12 meses, por lo tanto, resulta interesante analizar distintos espacios de tiempo para la desviación estándar con la que se calcula  $Y_t^i$ , por lo que se consideran además del propuesto por la publicación que es 1 año (252 días), los siguientes: 3 (63 días), 6 (126 días) y 9 (189 días) meses.

También, la publicación propone una función de posición que utiliza el indicador de tenencia. Esta función tiene como objetivo determinar la colocación óptima en cada clase de activo, pero dadas las restricciones del problema que resuelve el trabajo de título no es posible emplearla, sin embargo, esta señal es utilizada para generar el portafolio final.

La función de posición se define como:

$$X_t^i = \phi(Y_t^i)$$

Con:

$$\phi(y) = \frac{y * \exp\left(-\frac{y^2}{4}\right)}{0.89}$$

Lo llamativo de esta función viene de su forma, la cual se muestra a continuación:

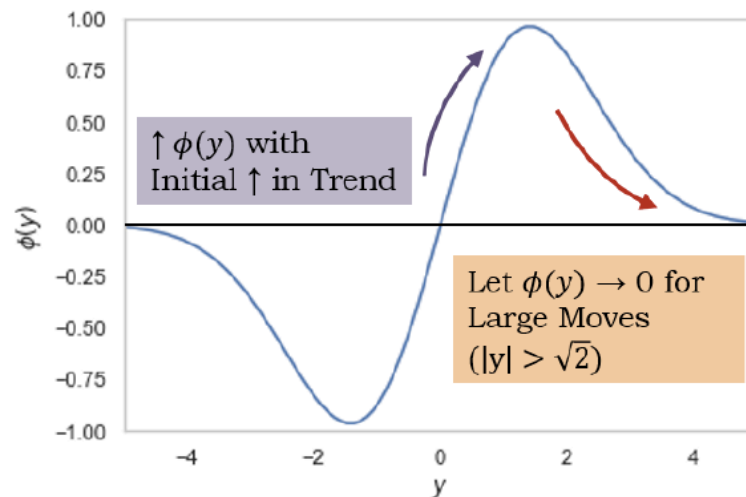


Figura 53: Función de Posición Momentum.

Como es posible apreciar en la figura, la función determina el tamaño de la posición en base a la tendencia del activo, en este sentido, asigna valores más grandes (pequeños) a medida que se espera que el activo vaya al alza (baja), es decir, cuando la tendencia está comenzando (ya sea de apreciación o depreciación) se toman posiciones más grandes que cuando ya lleva demasiado tiempo y se espera que cambie de sentido.

### 8.3.4. Construcción de Portafolio

Para la construcción del portafolio se utiliza la señal de posición como una puntuación. Dada las condiciones del problema, donde se tiene que generar un portafolio que toma posiciones largas y cortas que suman 0, se diseña la siguiente regla de posición, la cual ordena la puntuación de mayor a menor con el objetivo de tomar posiciones más grandes en aquellas clases de activo que se espera vayan a tener un rendimiento superior y pequeñas en aquellas que vayan a tener un rendimiento inferior, de la siguiente manera:

$$\bar{x} = [-\text{limite}, -\frac{\text{limite}}{2}, 0, \frac{\text{limite}}{2}, \text{limite}]$$

Luego, se debe ajustar la volatilidad del portafolio si es que sobrepasa a la definida por el mánager de inversión, por lo que el portafolio propuesto es multiplicado por un ponderador  $w$  tal que:

$$\sqrt{w * \bar{x} * VarCovar * \bar{x} * w} \leq TargetVol$$

Por último, se debe cumplir la restricción de que las posiciones en cada clase de activo no sean mayores al límite establecido por el mánager. Para resolver este problema se decide que, en caso de que la restricción no se cumpla:

- La diferencia entre la posición y el límite para renta fija local y renta variable emergente es cerrada con cash CLP.
- La diferencia entre la posición y el límite para renta fija extranjera, renta fija US y renta variable mercados desarrollados es cerrada con cash US.

También se debe cumplir que la suma de las clases de activo por tipo (renta variable y renta fija) no supere los límites establecidos por el mánager. Para resolver este problema en caso de que la restricción no se cumpla, se pondera la suma por tipo de activo por un valor  $z$  tal que:

$$\begin{aligned} |(x_{rve} + x_{rvd})| * z &\leq \text{limite}_{suma} \\ |(x_{rfl} + x_{rfhy} + x_{rfig})| * z &\leq \text{limite}_{suma} \end{aligned}$$

Luego:

- La diferencia entre la posición final y la posición inicial para renta fija local y renta variable emergente es cerrada con cash CLP.
- La diferencia entre la posición final y la posición inicial para renta fija extranjera, renta fija US y renta variable mercados desarrollados es cerrada con cash US.

### 8.3.5. Análisis y Backtest

Para el análisis de la estrategia, en un inicio, se utiliza el mismo periodo estudiado en los modelos KM y GMM. Esto principalmente por dos motivos, en primer lugar, poder comparar los modelos y, en segundo lugar, utilizar las mismas matrices de varianza-covarianza para ajustar el riesgo de la estrategia.

De los resultados obtenidos en KM y GMM se evidencia que una de las estrategias que mejor performance tiene y, a la vez, es interesante por la forma de la curva es rolling 48 meses. Por lo tanto, se muestra que tanto afecta la diferencia entre las matrices de varianza-covarianza calculadas en cada modelo en los resultados del modelo momentum.

Dicho esto, se analizan las tuplas propuestas por Lim, Zohren y Roberts (2019), junto a una cuarta señal que viene dada por el promedio de las señales correspondientes a las tuplas (8,24), (16,48) y (32,96), tal que  $\bar{Y}_t^i = \sum_k \frac{Y_t^i(S_k, L_k)}{3}$ .



Figura 54: Rentabilidad Acumulada Matriz Var-Covar GMM. [PB]

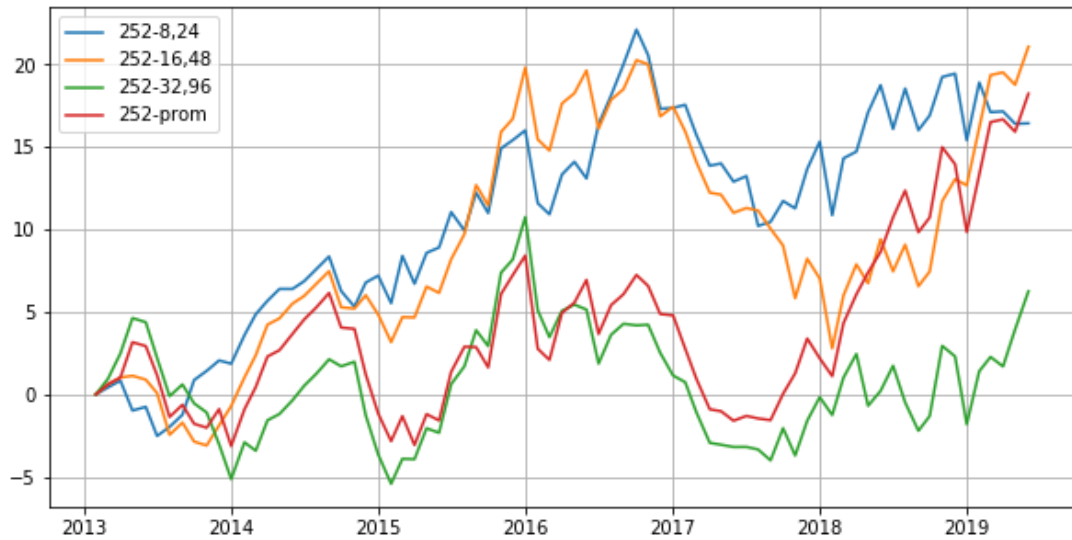


Figura 55: Rentabilidad Acumulada Matriz Var-Covar KM. [PB]

Como se evidencia en las figuras, la performance de la estrategia no se ve afectado por la matriz de varianza-covarianza que se utiliza para ajustar la volatilidad del portafolio, para confirmarlo, se comparan los portafolios asociados a las estrategias con mejor performance (252-16,48):

Index	DMEQLCEW Inde:	EMEQCULC Inde>	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	-15.1052	-2.13173	12.1189	-27.3161	32.4342
std	39.4342	42.2378	41.2915	63.0932	43.5217
min	-78.0547	-87.84	-100	-100	-100
25%	-50	-38.1783	-17.3687	-100	0
50%	-24.7263	0	0	-40.0347	27.9379
75%	0	30.9631	50	36.1541	54.6227
max	87.84	78.0547	100	100	100

Figura 56: Caracterización Portafolio Matriz Var-Covar GMM. [PB]

Index	DMEQLCEW Inde:	EMEQCULC Inde>	.CHFICUAA Index	LUATTRUU Index	HBFIWOCO Index
count	77	77	77	77	77
mean	-15.1667	-2.13849	12.1715	-27.2432	32.3768
std	39.0617	41.9607	41.2465	62.997	43.5371
min	-75.8645	-83.6273	-100	-100	-100
25%	-50	-38.2858	-17.4884	-99.6557	0
50%	-24.6566	0	0	-40.0806	27.8781
75%	0	31.3517	50	36.5222	54.3
max	83.6273	75.8645	100	100	100

Figura 57: Caracterización Portafolio Matriz Var-Covar KM. [PB]

Si bien existen diferencias en la descripción de cada portafolio, estas son mínimas del orden de 4 punto base (0.04%) con promedios y desviaciones estándar prácticamente iguales, por lo que se concluye que el modelo es indiferente a la diferencia de estas dos matrices. También hay que destacar el hecho de que no se cierran posiciones con Cash CLP o USD, lo que implica que nunca se supera que la suma por tipo de asset class sea mayor al límite<sup>64</sup>, esto es provocado por el ajuste que se hace en los pesos para la volatilidad.

Por otro lado, resulta interesante la performance de la estrategia. En primer lugar, se evidencia que no es basada en renta fija, en el sentido de que es posible apreciar volatilidad en la performance, implicando posiciones en renta variable. Esto se confirma al observar la caracterización de los portafolios hecha al inicio. También, y a diferencia de los modelos de ciclo económico, se tiene que independiente de la combinación el periodo correspondido entre 2017 y 2018 obtiene pérdidas, por consiguiente, se subentiende que en este periodo ocurrió algún acontecimiento tal que el modelo no es capaz de capturar.

Un punto relevante que viene a complementar lo expuesto en los modelos anteriores, es el hecho de que, por ejemplo, para la combinación 252-8,24 parte del periodo que está “flat” en los modelos de ciclo económico presenta rendimiento positivo en momentum, lo que genera un primer hallazgo puesto que estos modelos puedan ser complementarios.

A su vez, se debe analizar el hecho de que el mercado de renta fija local es pequeño y menos líquido que el resto, por lo que es más complejo realizar transacciones, por lo tanto, el mánager de inversión, tal y como para los modelos económicos, considera que un límite máximo de 25 puntos base (0,25%) es razonable. Obteniéndose los siguientes resultados:

<sup>64</sup> Dada los límites propuestos y la formulación del problema, es imposible que cada asset class supere el límite. El único caso posible es que la restricción de volatilidad cambie a igualdad, en este caso, si la volatilidad del portafolio per se es menor al objetivo, se aumentan los pesos tal que se alcance, implicando que sea posible que una clase de activos supere el límite.



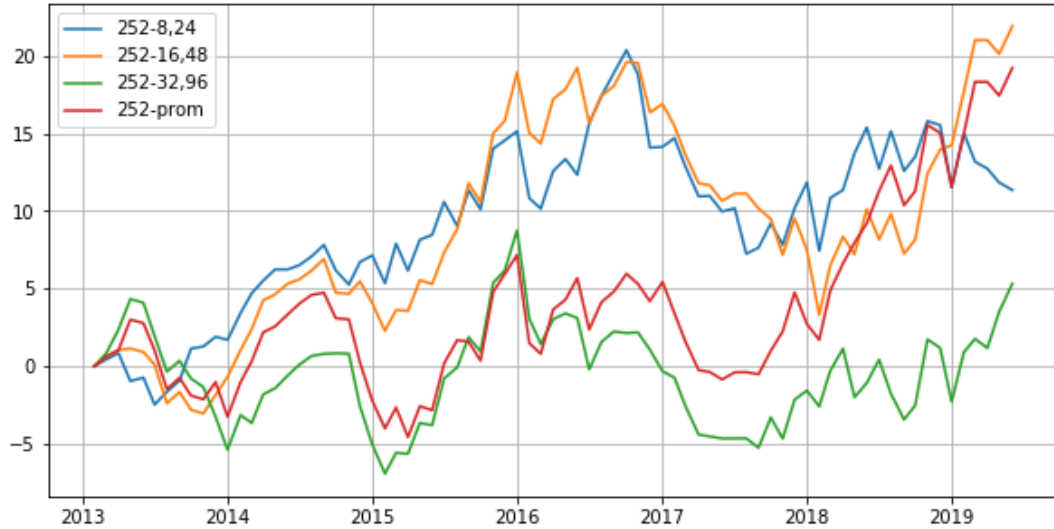


Figura 58: Rentabilidad Acumulada RFL Restringido. [PB]

Como es posible apreciar, la performance no se ve seriamente afectado, lo cual es un resultado esperado y no significa que el modelo sea mejor al de ciclo económico considerando esta restricción, puesto que no son comparables. El diseño de este modelo implica que el exceso de posición en renta fija local sea cubierto por cash CLP, por lo que es correcto señalar que, en general, cash CLP se comporta de manera similar a renta fija local, pero con menor retorno.

### 8.3.6. Análisis de Sensibilidad

#### Señal a 3, 6, 9 y 12 Meses

Para este caso se tienen 17 estrategias, pues como se ve en el análisis ya realizado, cada combinación aporta con 4 estrategias, además, se le agrega otra señal que corresponde al promedio de los promedios tal que:  $\hat{Y}_t^i = \frac{\sum_k \bar{Y}_k^i}{4}$ ;  $k \in \{3,6,9,12\}$

Se analiza el caso considerando como límite de inversión  $\pm 1\%$  por clase de activo y por suma de tipo de activo. Dada la gran cantidad de combinaciones, se filtraron en dos grupos, el primero muestra aquellos 6 con mejor rendimiento y, el segundo grupo, aquellos 6 con peor rendimiento. Los resultados son los siguientes:

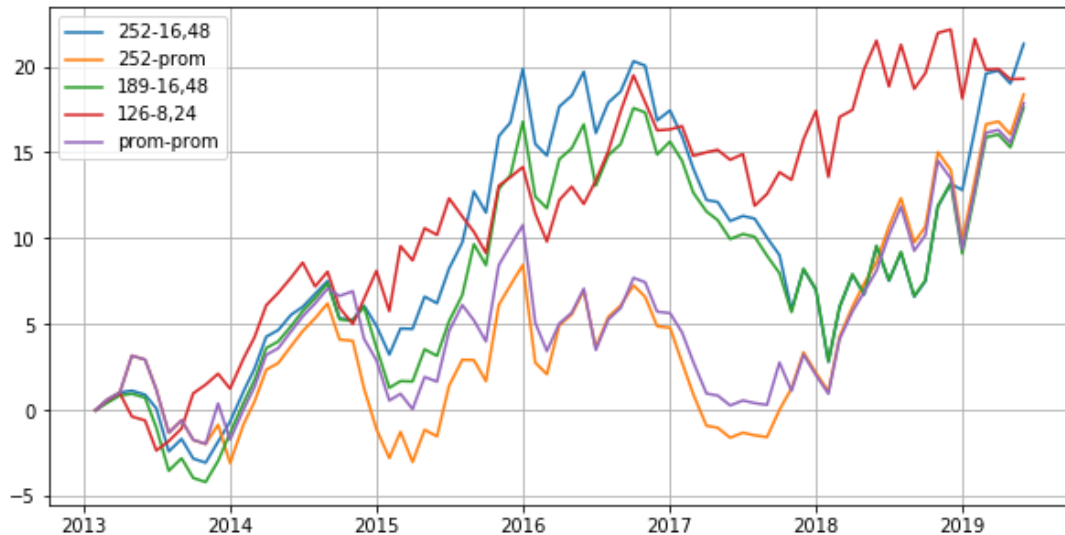


Figura 59: Rentabilidad Acumulada Mejores Combinaciones. [PB]

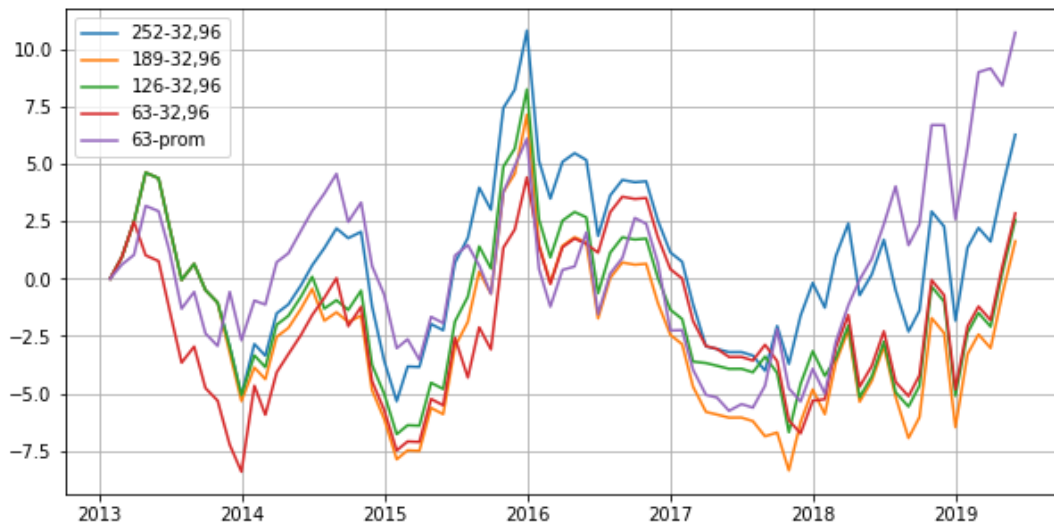


Figura 60: Rentabilidad Acumulada Peores Combinaciones. [PB]

Tal y como se aprecia de las figuras, existen ciertos patrones entre los modelos con mejor y peor rendimiento, por un lado, se tiene que aquellos que consideran menos días para genera MACD (implicando menor vida media) y, a su vez, mayor periodo para calcular la desviación estándar de activo, en general, poseen mejor performance que aquellos que consideran más días para generar MACD y menor periodo para el cálculo de la desviación estándar del activo. El poseer menor vida media implica que el decaimiento de la data es mucho más rápido, es decir, la data más antigua es menos relevante, por lo tanto, el desarrollar la señal MACD con menor vida media implica capturar las tendencias más cortoplacistas y, dada la estructura de las formulas, la desviación estándar de cada activo en 252 días (1 año) vendría a ser una medida de ajuste a la señal MACD, pues considera que tanto varía el activo en un periodo de tiempo mayor, siendo en sí un ajuste implícito de riesgo, lo que explica el resultado obtenido.

## Irrestricto en Volatilidad

Se realiza el análisis del modelo irrestricto en volatilidad con las mismas características utilizadas para los modelos de ciclo económico, que se resumen en límites de  $\pm 1\%$  en cada clase de activo y suma por tipo (renta variable y renta fija).

Dada la cantidad de combinaciones posibles, a continuación, se muestran las 5 con mejor y peor rendimiento:

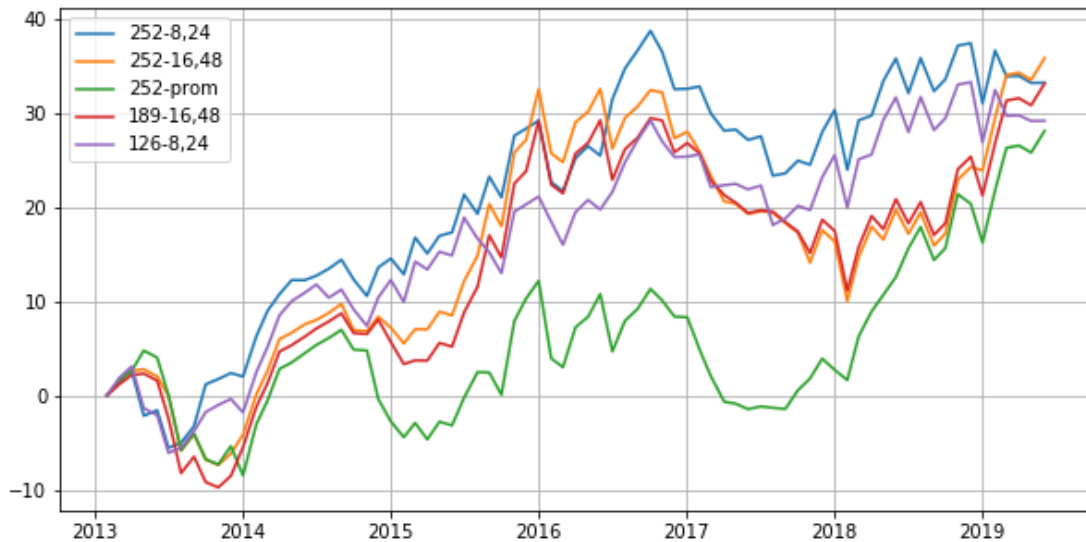


Figura 61: Rentabilidad Acumulada Mejores Combinaciones Estrategia Sin Vol-Obj. [PB]



Figura 62: Rentabilidad Acumulada Peores Combinaciones Estrategia Sin Vol-Obj. [PB]

Como es posible evidenciar, las características expuestas en el apartado anterior se repiten, lo cual es un resultado esperado, puesto que al no ajustar por volatilidad y, dada la construcción del portafolio, las posiciones que generan valor y pérdida son amplificadas y, en caso de superar el límite, cerradas por caja, que, a juicio de experto, en general poseen retorno positivo y pequeño.

También es relevante analizar que, si bien el modelo parece generar mayor rentabilidad en el periodo estudiado que los modelos de ciclo económico, esta no es suficiente por si sola, por lo que se debe medir la volatilidad de estos. Se tienen las siguientes volatilidades promedio para cada estrategia:

- Estrategias con buen performance:

<b>Volatilidad [PB]</b>	<b>252-8,24</b>	<b>252-16,48</b>	<b>252-prom</b>	<b>189-16,48</b>	<b>126-8,24</b>
<b>Promedio</b>	3.47	3.37	3.20	3.34	3.34

Tabla 17: Volatilidad Mejores 5 Estrategias Sin Vol-Obj. [PB]

- Estrategias con mal performance:

<b>Volatilidad [PB]</b>	<b>252-32,96</b>	<b>189-32,96</b>	<b>126-32,96</b>	<b>63-32,96</b>	<b>63-prom</b>
<b>Promedio</b>	3.12	3.10	3.12	3.20	3.13

Tabla 18: Volatilidad Peores 5 Estrategias Sin Vol-Obj. [PB]

Como se aprecia, las volatilidades promedio para cada estrategia no distan demasiado de las encontradas en los modelos de ciclo económico. Por lo que, si se ajusta la performance de los distintos modelos irrestricto en volatilidad al riesgo promedio, en primera instancia, momentum parece mejor en el periodo estudiado de 2013 hasta el 31 de mayo de 2019, sin embargo, se debe considerar el hecho de que las estrategias de ciclo económico no cierran posiciones con cash al estar irrestrictas, condición que en este modelo sí se presenta cuando la suma por tipo de activo es mayor al límite, por lo que no se puede asegurar que uno sea mejor que el otro pues no son del todo comparables.

Por otro lado, como se evidencia que cerrar posición con cash CLP cuando se restringe el límite en renta fija local no afecta en demasía la performance de las estrategias, no agrega valor analizar este tipo de portafolio.

## Periodo Completo

También resulta relevante analizar el rendimiento de la estrategia a lo largo de todo el periodo. Dado que para ajustar el modelo a una volatilidad objetivo el óptimo es utilizar la matriz de varianza-covarianza calculada a partir del modelo de ciclos económicos, pero por formulación de este no es posible calcular un buen valor para un periodo anterior al 2013, se realiza el análisis sin hacer el ajuste, por lo tanto, se enfoca en la forma de las curvas más que en el rendimiento de estas, puesto que no es posible calcular una buena medida de riesgo para la que valga la pena analizar el rendimiento.

Se comienza desde el 2004, puesto que tanto KM y GMM se encuentran en alta probabilidad en una fase, obteniéndose los siguientes resultados:

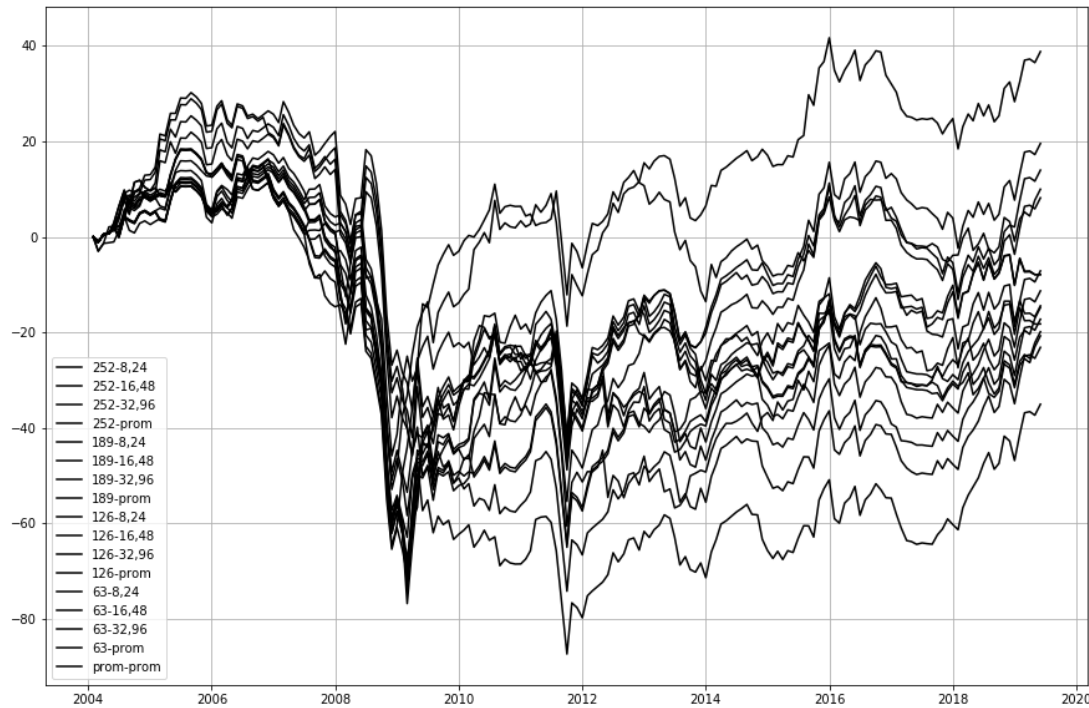


Figura 63: Rentabilidad Acumulada Periodo 2004-2019 Estrategia Sin Vol-Obj. [PB]

A nivel de masa, es claro evidenciar que existen distintos periodos interesantes a analizar, de los que se destacan:

- 2008-2009: Periodo en el que la performance de las estrategias fue consistentemente muy mala.
- 2010-2011: Periodo en el que la performance es relativamente plana.
- 2013-2019: Periodo en el que la performance de la estrategia parece estabilizarse, el cual es estudiado en este trabajo.

Que en un periodo largo se destaque una estrategia con respecto al resto no implica que sea consistentemente mejor, puesto que su performance puede venir dado a que en una parte acotada de tiempo la estrategia fue muy exitosa, generando gran diferenciación con respecto a las demás y, luego, se comporte igual o peor que las otras (como se observa en la figura). Esta es una de las justificaciones del porqué realizar análisis por periodo. Por otro lado, no se tiene una medida de, por ejemplo, riesgo con la cual se puedan comparar las distintas estrategias, puesto que el mayor performance puede venir implícito a un mayor riesgo.

## 2008-2009



Figura 64: Rentabilidad Acumulada Periodo 2008-2009 Estrategia Sin Vol-Obj. [PB]

Este periodo es uno de los más llamativo a analizar, en primer lugar, por el pobre performance del modelo, el cual es consistentemente malo para cada una de las estrategias y, en segundo lugar, que el pobre performance ocurre en el mismo periodo de la crisis sub-prime<sup>65</sup>, una de las crisis económicas mundiales más grandes del último tiempo<sup>66</sup>.

<sup>65</sup> Wikipedia. Crisis Financiera de 2008. Disponible en: [https://es.wikipedia.org/wiki/Crisis\\_financiera\\_de\\_2008](https://es.wikipedia.org/wiki/Crisis_financiera_de_2008) (consulta: 11 de enero de 2020).

<sup>66</sup> Finance Yahoo. Las Crisis Económicas Más Importantes del Mundo. Disponible en: <https://tinyurl.com/vg9kunt> (Consulta: 11 de enero de 2020).

Considerando la correlación entre el abrumador performance negativo del modelo y la crisis sub-prime y, además, la consistencia del modelo desde 2013 en adelante, donde la pendiente de las estrategias del modelo tiende a ser positiva, surge la incógnita de como lograr que el modelo se ajuste en periodos de crisis, para lo cual se tiene que si se observa la evolución de las probabilidades de KM y GMM, es evidente que el modelo es capaz de diferenciar este periodo del resto, por lo que en el caso de que el resultado de estos sea favorable, generar un mix es altamente conveniente, logrando así complementarse.

## 2010-2011

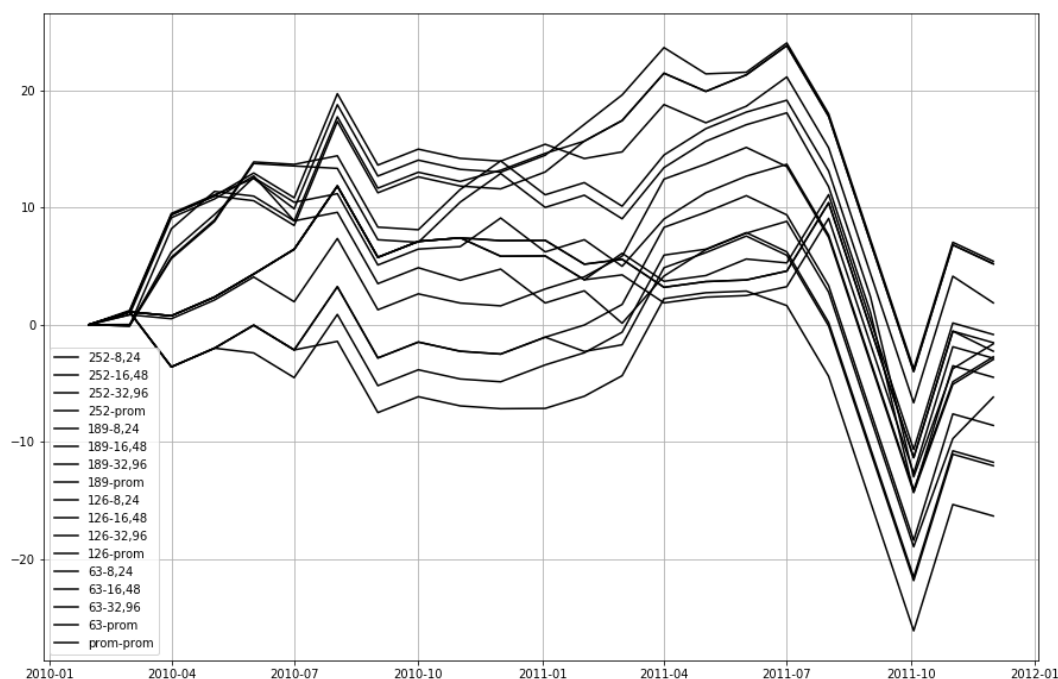


Figura 65: Rentabilidad Acumulada Periodo 2010-2011 Estrategia Sin Vol-Obj. [PB]

Este periodo viene luego de la recesión, en el cuál es posible observar que de extremo a extremo las distintas estrategias del modelo no son capaces de generar rentabilidad, lo cuál es un resultado esperado pues, dado que la crisis es reciente, se puede catalogar a este transcurso de tiempo como de transición, en el cual hay incertidumbre. Este hecho complementa lo expuesto en el punto anterior, que es la búsqueda de modelos complementarios capaces de que, en la mezcla, obtengan curvas persistentes y consistentes.

### 8.3.7. Síntesis - Momentum

En base a los resultados obtenidos, las distintas estrategias de momentum poseen un perfil atractivo para el mánager, pues se evidencia que toman posiciones tanto en activos de renta fija y renta variable, sin embargo, la performance no es tal para que genere diferenciación, más aún, si se considera que en el caso restricto se obtienen retornos acumulados para el periodo inicial estudiado peores que los asociados a las estrategias de los modelos de ciclo económico, por lo que, en primera instancia, no se cumple el objetivo principal del trabajo de título.

Por otro lado, se tiene que, en general, considerar un decaimiento más rápido de la data para generar la señal implica mejores retornos que considerar un periodo más largo, por lo tanto, la señal es capaz de capturar la tendencia y el impulso de buena manera en el corto plazo.

Complementando el párrafo anterior, se tiene que la estrategia no parece funcionar bien ante periodos de crisis (en los que existe alta incertidumbre), lo que puede explicarse que, al tener un perfil de corto plazo, la volatilidad existente en el periodo dada la incertidumbre provoca que las señales a corto plazo no sirvan de indicadores, puesto que los cambios son constantes y grandes.

Por otro lado, un resultado destacable es el encontrado al no considerar una volatilidad objetivo, puesto que las distintas estrategias calculadas poseen mejores retornos acumulados en el periodo correspondido entre 2013 y mayo de 2019 que sus símiles del modelo de ciclo económico, más aún, si se considera que las estrategias, en general, presentan performance positivo cuando las estrategias de los modelos de ciclo económico parecen estar “flat”. Por lo tanto, si bien la estrategia por si sola no parece cumplir con el objetivo principal del trabajo de título, abre la posibilidad a ser utilizada como señal de posición cortoplacista para una mezcla.



#### 8.4. Mezcla de Modelos

Según los resultados obtenidos, resulta interesante analizar la mezcla de los distintos modelos, esto con el objetivo de confirmar algunos hallazgos visuales que vienen a ser un aporte considerable al trabajo de título, especialmente por el hecho que históricamente se ha postulado la diversificación de riesgo a través de distintos activos tal como expone Leavens (1945), por lo tanto, este trabajo analiza la segunda derivada de esta postulación, que se define como diversificar riesgo a través de estrategias.

Comúnmente se emplea la analogía de diversificar riesgo en distintos activos como no colocar todos los huevos en una misma canasta. Para esta segunda derivada se puede emplear la analogía de que no todas las canastas las lleve el mismo repartidor.

Por otro lado, considerando las múltiples estrategias que se pueden desarrollar en cada modelo, implicando una cantidad aún mayor de combinaciones, se decide a juicio de experto que lo más razonable es combinar los modelos de la manera menos restricta posible, para luego ajustarlos a las restricciones estipuladas, puesto que restringir una mezcla de modelos restringidos implica una especie de doble marginalización, lo que puede provocar pérdida de rentabilidad.

Con esto en consideración, se toman estrategias restrictas solo por el tamaño límite por clase de activo para cada modelo que, a juicio de experto, es 1% cuando existe estabilidad en el mercado:

- Ciclo económico KM: Sin volatilidad objetivo ni restricción de límite por tipo de activo (renta fija y renta variable).
- Ciclo económico GMM: Sin volatilidad objetivo ni restricción de límite por tipo de activo (renta fija y renta variable).
- Momentum: Sin volatilidad objetivo ni restricción de límite por tipo de activo (renta fija y renta variable). Se escoge la con mejor performance, correspondiente a la calculada a partir de la tupla (8,24) a 1 año.

Obteniéndose las siguientes combinaciones:

- KM-Momentum.
- GMM-Momentum.
- KM-GMM.
- KM-GMM-Momentum.

Estos portafolios al ser la mezcla igualmente ponderada de cada modelo se deben ajustar al riesgo y las restricciones de posición. Con esto en consideración, es importante destacar que, a juicio de experto, cuando los límites de posición corresponden a 1% se tiene una volatilidad objetivo de 2 puntos base.

En primer lugar, las posiciones para cada fecha se ajustan a la volatilidad objetivo, esto considerando un  $w$  que pondera las posiciones tal que:

$$\sqrt{w * \bar{x} * VarCovar * \bar{x} * w} = TargetVol$$

Es importante destacar el hecho de que si la mezcla es GMM-MOM se utiliza la matriz de varianza-covarianza de GMM, a su vez, si es KM-MOM, se utiliza la matriz de varianza-covarianza de KM, por último, si es del tipo KM-GMM se utiliza el promedio de las matrices de varianza-covarianza.

Luego, las posiciones se ajustan en base a los límites establecidos tanto para cada clase de activo como para tipo de activo que considere el mánager. En este sentido, se establecen los siguientes límites para cada uno de los casos:

	%
<b>Suma Renta Variable</b>	1%
<b>Suma Renta fija</b>	1%
<b>RVD</b>	1%
<b>RVE</b>	1%
<b>RFHY</b>	1%
<b>RFIG</b>	1%
<b>RFL</b>	0.25%

Tabla 19: Restricciones de Posición.

Con esto en consideración, se ajustan los límites tal y como se hace para el modelo momentum.

### 8.4.1. Correlación

Para calcular la correlación de las estrategias se utilizan los retornos mensuales obtenidos por cada modelo. Cabe destacar que la razón de calcular la correlación a partir de los retornos mensuales viene de la siguiente lógica: si dos modelos están altamente correlacionados se espera que ambos obtengan retornos positivos (negativos) en el mismo periodo por cada posición que toman. Obteniéndose los siguientes resultados:

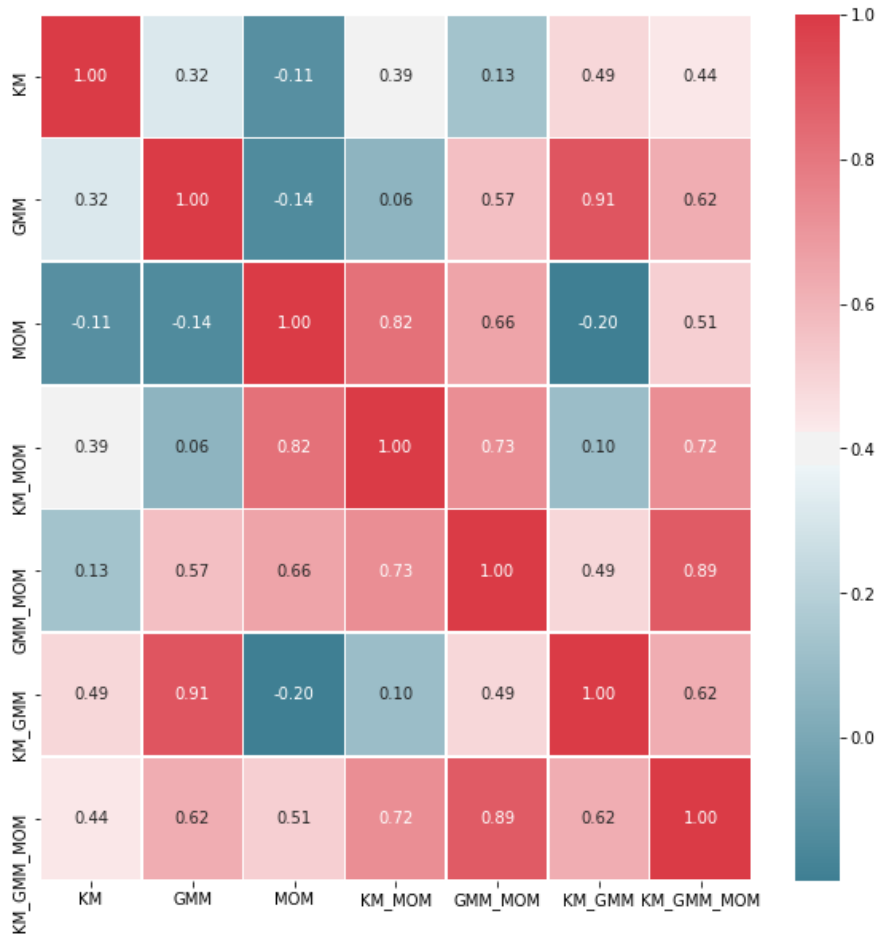


Figura 66: Correlación Modelos.

Uno de los principales hallazgos encontrados es que las estrategias de ciclo económico presentan una correlación cercana a 0 con la de momentum. Un resultado bueno para el objetivo principal del trabajo de título, puesto que implica que la performance de estos no se mueve de la misma forma, por lo tanto, parte de los periodos que se evidencian como “flat” en los modelos de ciclo económico tienen retorno en el de momentum y viceversa. Por consiguiente, generar un mix de estos modelos implica obtener retorno positivo a largo plazo.

Otro hallazgo relevante es la correlación entre los modelos de ciclo económico, estos poseen las mismas características en términos de diseño, diferenciándose solamente por la forma en que se calculan las probabilidades, que en sí implica bastante, puesto que en base a las probabilidades se calculan las matrices de varianza-covarianza, retornos esperados por fase y portafolios. Por lo tanto, se confirma la sensibilidad del modelo a esta variable y, por consiguiente, la relevancia de la data utilizada en el modelo tal que calidad de esta sea capaz de diferenciar las fases y calcular las probabilidades de buena manera.

Por otro lado, se tiene que las mezclas de modelos variados no poseen una correlación suficientemente alta ( $>0.8$ ) como para asumir que la performance de estos se mueve conjunta, por lo que cada estrategia mezclada genera valor de distinta manera.

Por último, surge la incógnita del porqué de las correlaciones encontradas. La cual puede responderse, principalmente, por la construcción de los modelos, por un lado, el modelo momentum genera su posición considerando variables a corto plazo, como la vida media a 8 y 24 días con un ajuste de desviación estándar a 12 meses. Por otro lado, los modelos de ciclo económico generan sus posiciones considerando aspectos a más largo plazo, como el retorno esperado a 6 meses y matrices de varianza-covarianza a 48 meses. Por consiguiente, estas características pueden explicar la razón de que los modelos sean complementarios.

#### 8.4.2. Performance y Métricas

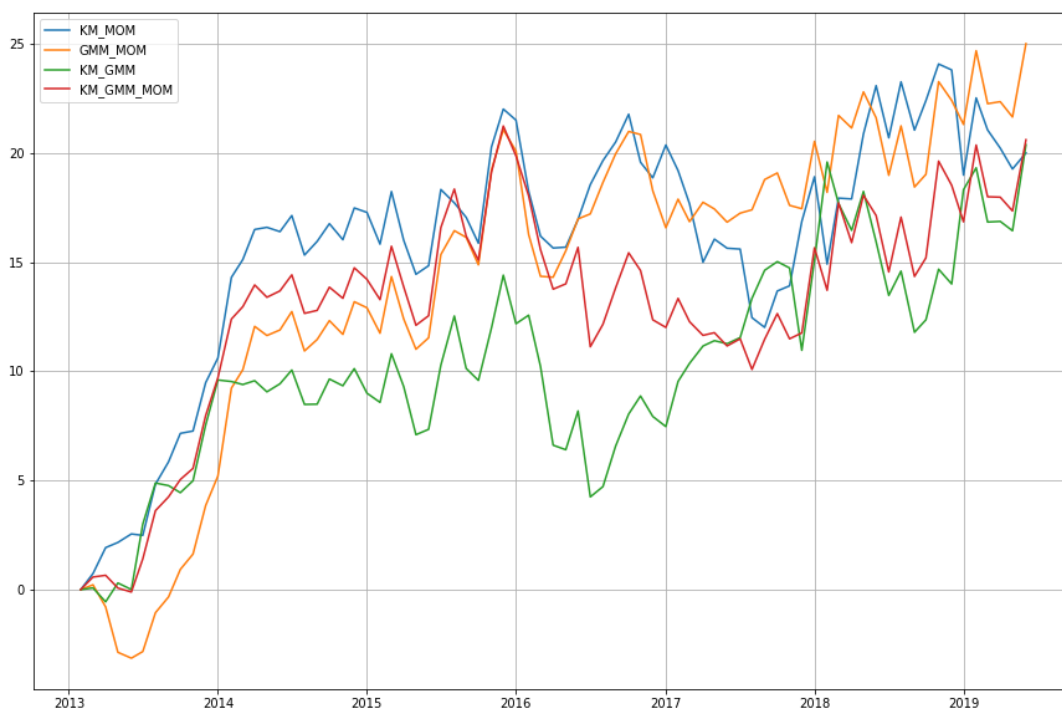


Figura 67: Rentabilidad Acumulada Mezcla de Modelos. [PB]

Como es posible apreciar, a lo largo del periodo estudiado una de las mejores estrategias en términos de rentabilidad acumulada y forma de la curva es la mezcla entre momentum y GMM, resultado que también concuerda con las correlaciones encontradas en el apartado anterior y la performance de las estrategias encontradas cuando se analizan los modelos por separado.

Por otro lado, llama la atención la performance y forma de la mezcla entre KM y momentum, sin embargo, esta se explica por la forma de las curvas de los modelos irrestrictos que se detallan a continuación:

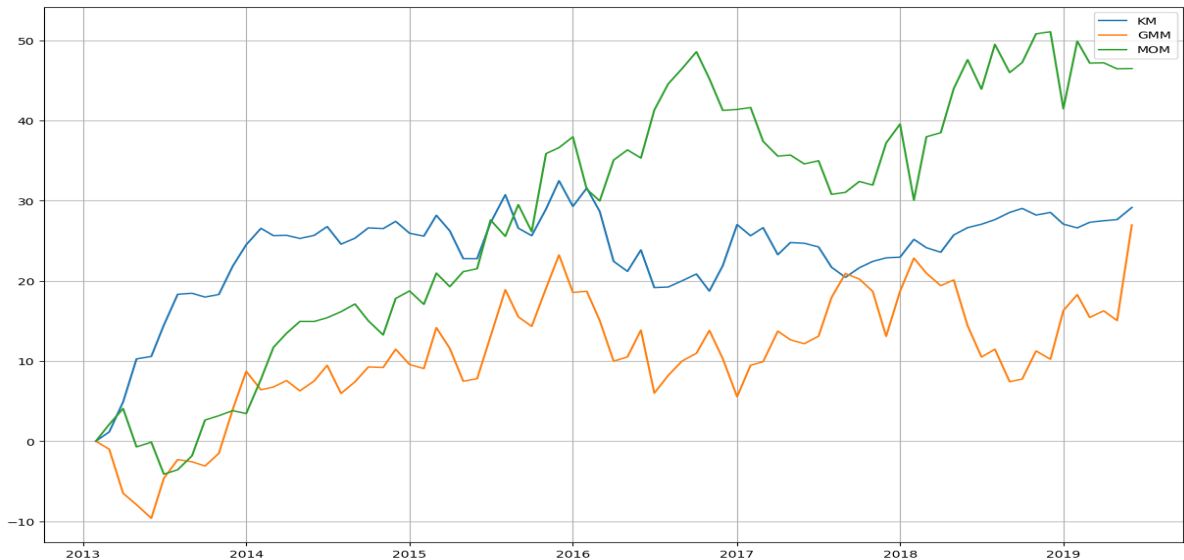


Figura 68: Rentabilidad Acumulada Modelos Irrestrictos en Volatilidad. [PB]<sup>67</sup>

Como es posible apreciar, la curva KM parece estar flat la mayor parte del tiempo desde 2014 en adelante, lo que implica que la performance de la mezcla no tenga la misma pendiente que GMM-Momentum, sin embargo, el desempeño de la estrategia KMM de 2013 a 2014 y de momentum entre 2018 y 2019 es sobresaliente, lo que explica aquellos saltos en la rentabilidad de KM-MOM.

Por otro lado, se observa que entre menos correlación se tenga entre modelos, la performance total de la mezcla parece ir mejor, lo cual es un resultado esperado pues una correlación cercana a 0 implica que cuando uno de los modelos se mueve, el otro no, por lo tanto, la mezcla captura el movimiento de ambos.

<sup>67</sup> Es importante recalcar el hecho de que la performance de estas estrategias no es comparable con ninguna otra analizada hasta el momento, puesto que corresponde a los modelos optimizados sin volatilidad objetivo ni restricción de suma por tipo de clase de activo (renta fija y renta variable), esta segunda restricción siempre fue activa para todos los análisis anteriores y se decidió quitar, a juicio de experto, para mezclar las performances lo más puras posibles de los modelos.

También resulta relevante analizar métricas de los distintos portafolios calculados, las que se detallan a continuación:

- CAGR: Corresponde al retorno de la estrategia a lo largo de todo el periodo anualizado.  $CAGR = \frac{r_a}{años}$ .
- Volatilidad Anual: Corresponde al promedio de las volatilidades anuales de cada posición por estrategia.  $Anual\ Vol = \sum_i \frac{Vol-Anual_i}{I}$ .
- Ratio: Es una medida de la compañía y corresponde al ratio entre CAGR y Volatilidad anual. Se puede interpretar como que porcentaje de la volatilidad obtuve en un año.  $Ratio = \frac{CAGR}{Volatilidad-Anual}$ .
- Max DD: Corresponde a la máxima baja en términos de retorno acumulado de la estrategia.  $Max\ DD = Max\ R_a - Min\ R_a$ , tal que el mínimo ocurra después del máximo.
- Hit Ratio: Corresponde a la razón en la que la estrategia tuvo retornos positivos con respecto a la cantidad total de posiciones que tomó.  $Hit\ Ratio = \frac{\sum Retorno\ mensual(>0)}{\#Transacciones}$ .
- Curtosis: Corresponde a una medida para ver que tan distribuidos cerca de la media se encuentran los retornos mensuales. En este caso se mide el exceso de curtosis con respecto a una distribución normal, es decir, curtosis = 0 implica distribución normal.
- Skewness: Corresponde a una medida para ver la simetría de los retornos mensuales obtenidos. Se tiene que Skewness = 0 implica distribución simétrica.

Para las mezclas de modelos se obtienen los siguientes resultados:

	KM_MOM	GMM_MOM	KM_GMM	KM_GMM_MOM
<b>CAGR [PB]</b>	3.12	3.89	3.17	3.21
<b>Anual Vol. [PB]</b>	6.65	6.77	6.92	6.77
<b>Ratio</b>	0.47	0.58	0.46	0.47
<b>Max DD [PB]</b>	5.09	0.00	0.00	11.15
<b>Hit Ratio</b>	0.57	0.57	0.57	0.55
<b>Curtosis</b>	0.05	-0.18	-0.02	-0.29
<b>Skewness</b>	-0.26	0.22	-0.02	0.19

Tabla 20: Métricas de Medición.

Como es posible observar de la tabla, el modelo GMM-MOM es el que posee mejores indicadores a nivel agregado del periodo a estudiar, puesto que tal y como se observa en términos de rentabilidad, genera retorno consistentemente.

Otro aspecto relevante es analizar la distribución de los retornos mensuales de las estrategias. Una curtosis negativa implica que los retornos están menos concentrados en la media (se adjunta figura en anexo 5), lo que se complementa con la skewness, si esta es negativa implica que la media es menor que la mediana y esta, a su vez, es menor que la moda (Lo contrario en caso positivo. Se adjunta figura en anexo 6), por lo que la cola asociada a eventos de la extrema izquierda es más pesada (son más grandes), pero menos usual. Si bien los valores son pequeños en sí, implicando que los efectos de las distintas características de curtosis y skewness no sean abrumadores, para cada estrategia se tiene:

- **KM-MOM:** Datos más concentrados en la media que una normal, con una distribución cargada a eventos positivos de retorno y con una cola izquierda más pesada, es decir, cuando un retorno es negativo, es probable que sea muy negativo con respecto a lo usual.
- **GMM-MOM:** Datos menos centrados en la media, con una distribución cargada hacia retornos menores a la media, con una cola derecha más pesada, es decir, cuando un retorno es positivo, es probable que sea muy positivo con respecto a lo usual.
- **KM-GMM:** Datos menos centrados en la media, con una distribución cargada hacia retornos menores a la media, con una cola izquierda más pesada, es decir, cuando un retorno es negativo, es probable que sea muy negativo con respecto a lo usual.

- **KM-GMM-MOM:** Datos menos centrados en la media, con una distribución cargada hacia retornos menores a la media, con una cola derecha más pesada, es decir, cuando un retorno es positivo, es probable que sea muy positivo con respecto a lo usual.

Estos resultados concuerdan con la performance encontrada en los modelos, implicando que, por ejemplo, KM-GMM sea la peor combinación de todas a nivel agregado apreciándose que, en distintos periodos, otras combinaciones hayan tenido mejor resultado.

Por otro lado, se tiene que el hit ratio de todas las estrategias son casi idénticos, por lo que la diferencia que se genera en la performance a nivel agregado de las estrategias viene dada, principalmente, por dos motivos, cuando una gana lo hace en mayor medida que las otras y/o cuando pierde lo hace en menor medida.

Por último, resulta interesante desagregar la performance por año. Comprendiendo el periodo de 2013 a 2018, 2019 queda fuera del análisis puesto que la data utilizada es hasta mayo de 2019, faltando más de la mitad del año, implicando que las métricas no sean comparables al resto. Se obtienen los siguientes ratios:

<b>Ratio</b>	<b>KM_MOM</b>	<b>GMM_MOM</b>	<b>KM_GMM</b>	<b>KM_GMM_MOM</b>
<b>2013</b>	1.36	0.56	1.09	1.17
<b>2014</b>	1.25	1.36	0.37	0.96
<b>2015</b>	0.69	1.17	0.60	0.91
<b>2016</b>	-0.45	-0.41	-0.92	-1.27
<b>2017</b>	-0.30	-0.12	0.45	-0.10
<b>2018</b>	1.01	0.75	0.44	1.02
<b>Promedio</b>	0.59	0.55	0.34	0.45

Tabla 21: Métricas de Medición Desagregado por Año.

Como se aprecia en la tabla, no hay un modelo que sea mejor a los demás en todo periodo, más aún, KM-MOM resulta ser el que, en promedio, es mejor al resto, puesto que en algunos periodos fue considerablemente bueno, rentando con creces más que la volatilidad anual del año. Por lo tanto, no es suficiente analizar la performance en un periodo agregado para las estrategias para determinar cual modelo es mejor, puesto que dadas las condiciones del mercado los modelos se pueden ver afectados de distinta manera y, en ese sentido, es donde la mezcla de modelos viene a agregar el mayor valor.



### 8.4.3. Síntesis – Mezcla de Modelos

En primer lugar, hay que destacar que el mayor hallazgo de esta sección y, también, del trabajo de título, viene de la ortogonalidad (baja correlación) de los modelos diseñados para generar rentabilidad para la compañía. Si bien la performance de las mezclas realizadas dista del valor para considerarse buenas al ajustarse por riesgo (44 puntos base), se genera un espacio para el desarrollo y mejoras de los distintos modelos, puesto que, tal y como se observa en el análisis de cada modelo, estos son sensibles a las restricciones que se le imponen y las variables a utilizar.

Por otro lado, y complementando el párrafo anterior, se tiene que el objetivo del trabajo de título es aumentar la rentabilidad de la compañía en base a modelos estadísticos/matemáticos. Considerando que las posiciones de los modelos que se mezclan son riesgosas y estas, luego de mezclarse, son ajustados a un nivel de riesgo tal que se obtiene un performance como el de GMM-MOM que, si bien por sí solo no genera la rentabilidad deseada, pero dado su diseño son atractivos, tanto que a juicio de experto se pueden utilizarse como directriz para tomar decisiones.

A su vez, se tiene que independiente de la forma en la que se mide el desempeño de las mezclas de los modelos, ya sea a nivel agregado, desagregado o utilizando distintas métricas, aquellas mezclas de modelos menos correlacionados presentan mejor rendimiento que el resto.

Por último, hay que destacar la relevancia de las métricas al momento de comparar modelos, más aún si se hacen mezclas de estos, puesto que hay periodos en los que distintas combinaciones andan mejor que otras y, por lo tanto, es posible generar mejores combinaciones que, por ejemplo, pueden ser condicionales del tipo: si se está en fase "X", se mezcla 30% de GMM, 60% momentum y 10% de KM. Este análisis queda propuesto para otra investigación.

## 9. CONCLUSIÓN

En primer lugar, se tiene que los distintos modelos generan rentabilidad, sin embargo, esta es, en promedio, un 65% del valor para que sea atractivo implementarlos, además, las soluciones encontradas son bastante sensible a cambios en cada una de las entradas que posee, como lo es la matriz de varianza-covarianza. Por lo tanto, queda propuesto utilizar otro método de optimización para generar el allocation final, el cual puede ser jerárquico para así diferenciar de mejor manera los activos de renta variable con respecto a los de renta fija (más y menos riesgosos).

También es relevante destacar el hecho de que los modelos son sumamente sensibles a las restricciones que se le imponga, existiendo casos en el que al restringir más se obtienen mejores resultados, dado que, por ejemplo, para los modelos económicos los retornos esperados por fase no están definidos de manera óptima, existiendo de esta manera un espacio de mejora, por lo que un modelo mezclado puede ser una excelente opción.

Complementando el párrafo anterior, es sumamente complejo generar una función que sea capaz de mapear la relación entre variables independiente con el retorno para una clase de activos cualquiera, principalmente por el hecho de que en el mundo financiero constantemente van ocurriendo cosas nuevas, por lo que es evidente que una función que aprende del pasado no será capaz de capturar ese comportamiento. Por lo que queda propuesto idear algún otro método que realice el mapeo de mejor forma.

A su vez, con la data utilizada fue posible capturar 4 fases económicas bien definidas para el modelo KM y 5 fases económicas bien definidas para el modelo GMM, con esto en consideración, queda propuesto agregar más variables con el objetivo de caracterizarlas de mejor manera y, por qué no, generar mejores resultados, puesto que no hay una medida exacta de que tanto de los movimientos del mercado explican las variables seleccionadas.

Por otro lado, el modelo momentum parece funcionar de buena manera en periodos estables y de baja incertidumbre, por lo tanto, desarrollar un modelo complementario del tipo sentimiento de riesgo capaz de capturar la incertidumbre del mercado puede potenciar y estabilizar los resultados obtenidos.

Si bien el objetivo principal del trabajo de título es generar exceso de retorno a través de modelos matemáticos/estadísticos, el cual no se cumple de buenas a primeras en el periodo estudiado entre 2013 y 2019, se destaca un hallazgo relevante que es la baja correlación entre los modelos realizados, resultado obtenido debido, principalmente, por las características de estos al considerar distintos periodos de tiempo para generar allocation. Dicho lo anterior, los modelos mezclados tienden a generar retorno positivo el 60% de las veces y, en general, poseen una pendiente positiva en el retorno acumulado, por lo tanto, queda propuesto para otra investigación el desarrollo de mezclas con distintos pesos y no igualmente ponderadas, las cuales maximicen las ganancias, minimicen las pérdidas y/o tengan mejor ratio de acierto.

Por último, se evidencia bajo simulaciones que los modelos de ciclo económico generan exceso de retorno tal que son posibles de implementar para otros periodos de tiempo que por formulación no son factibles de analizar como para 2013-2019, más aún, el modelo GMM parece ser una buena directriz de posiciones dada su características de probabilidades cercanas a 1 en cada fase, puesto que se evidencia que el retorno acumulado en periodos de 1 año para distintas fechas es, en promedio, 1.14 veces el esperado en el caso restringido. Por lo tanto, es interesante analizar este problema tiempo después, cuando no se tenga un periodo largo de tiempo concentrado en fases particulares.

## 10. BIBLIOGRAFÍA

- Assnes, C., Moskowitz, T. and Pederson, L. Value and Momentum everywhere. *The Journal of Finance*, 2013.
- Baz, J., Granger, N., Harvey, C. Le Roux, N. and S. Rattray. *Dissecting investment strategies in the cross section and time series*. 2015.
- Bourgeron, T., Lezimi, E., Roncalli, T. *Robust Asset Allocation for Robo-Advisors*. Amundi, 2018.
- Defend, M., Portelli, L. and Cesarini, F. *Advanced Investment Phazer*. Amundi, 2017.
- Defend, M., Portelli, L. and Cesarini, F. *Asset Class Assessment (ACA)*. Amundi, 2017.
- Defend, M., Portelli, L. and Cesarini, F. *Asset Class Quant Models*. Amundi, 2019.
- Husnain, M., Hassan, A., Lamarque, E. *Shrinking the Variance-Covariance Matrix: Simpler is Better*. *The Lahore Journal Economics*, 2016.
- Leavens, D. *Diversification of investments*. 1945.
- Lim, B., Roberts, S., Zohren, S. *Enhancing Time Series Momentum Strategies Using Deep Neuronal Networks*. 2019.
- Lopez de Prado, M. *A Robust Estimator of The Efficient Frontier*. 2019.
- *Post-Moderm Portfolio Construction*. JP Morgan, 2017.

## 11. ANEXOS

### 11.1. Anexo 1

## HARD CLUSTERING MODELS

Taken from the slides of the edX course ColumbiaX: CSMM.102x Machine Learning

### Review: K-means clustering algorithm

---

**Given:** Data  $x_1, \dots, x_n$ , where  $x \in \mathbb{R}^d$

**Goal:** Minimize  $\mathcal{L} = \sum_{i=1}^n \sum_{k=1}^K \mathbb{1}\{c_i = k\} \|x_i - \mu_k\|^2$ .

► Iterate until values no longer changing

1. Update  $c$ : For each  $i$ , set  $c_i = \arg \min_k \|x_i - \mu_k\|^2$
2. Update  $\mu$ : For each  $k$ , set  $\mu_k = (\sum_i x_i \mathbb{1}\{c_i = k\}) / (\sum_i \mathbb{1}\{c_i = k\})$

Figura 69: K-Means<sup>68</sup>.

---

<sup>68</sup> Sandipanweb. Hard & Soft Clustering. Disponible en: <https://sandipanweb.wordpress.com/2017/03/19/hard-soft-clustering-with-k-means-weighted-k-means-and-gmm-em/> (consulta: 31 de diciembre de 2019).

## EM FOR THE GMM

Taken from the slides of the edX course ColumbiaX: CSMM.102x Machine Learning

### Algorithm: Maximum likelihood EM for the GMM

**Given:**  $x_1, \dots, x_n$  where  $x \in \mathbb{R}^d$

**Goal:** Maximize  $\mathcal{L} = \sum_{i=1}^n \ln p(x_i | \pi, \mu, \Sigma)$ .

► Iterate until incremental improvement to  $\mathcal{L}$  is “small”

1. **E-step:** For  $i = 1, \dots, n$ , set

$$\phi_i(k) = \frac{\pi_k N(x_i | \mu_k, \Sigma_k)}{\sum_j \pi_j N(x_i | \mu_j, \Sigma_j)}, \quad \text{for } k = 1, \dots, K$$

2. **M-step:** For  $k = 1, \dots, K$ , define  $n_k = \sum_{i=1}^n \phi_i(k)$  and update the values

$$\pi_k = \frac{n_k}{n}, \quad \mu_k = \frac{1}{n_k} \sum_{i=1}^n \phi_i(k) x_i, \quad \Sigma_k = \frac{1}{n_k} \sum_{i=1}^n \phi_i(k) (x_i - \mu_k)(x_i - \mu_k)^T$$

Comment: The updated value for  $\mu_k$  is used when updating  $\Sigma_k$ .

Figura 70: GMM <sup>69</sup>.

El anexo 1 corresponde a la formulación matemática de, en primer lugar, el algoritmo de clustering K-Means y, en segundo lugar, el algoritmo de maximización de expectativas que se utiliza para GMM.

<sup>69</sup> Sandipanweb. Hard & Soft Clustering. Disponible en: <https://sandipanweb.wordpress.com/2017/03/19/hard-soft-clustering-with-k-means-weighted-k-means-and-gmm-em/> (consulta: 31 de diciembre de 2019).

## 11.2. Anexo 2

	CPI YOY Index	PMI Manu	Earnings Pasados	Sales Pasados	Past Earnings	Past Sales growth	Margin Pasado	TPM Fed Hoy	Deuda Total/	Deuda Neta/	Deuda Total/	Deuda Neta per Share
CPI YOY Index	1	0.104002694	-0.140286812	-0.2782394	0.11178232	0.27971965	0.12156316	0.54546243	0.45549795	0.40319571	0.41663056	0.519258732
PMI Manu	0.104002694	1	0.358762013	0.22205771	0.32286339	0.1857272	0.48885055	-0.163425	-0.2231	-0.1840048	-0.1988233	-0.146388705
Earnings Pasados	-0.140286812	0.358762013	1	0.94005014	0.07124254	0.10595012	0.86283003	-0.2835687	-0.7663804	-0.7597448	-0.7245236	-0.432468213
Sales Pasados	-0.278239437	0.222057714	0.940050141	1	0.01070741	0.051077	0.65182788	-0.4473412	-0.8650052	-0.8640854	-0.8256587	-0.549163104
Past Earnings growth rate	0.111782319	0.322863386	0.071242544	0.01070741	1	0.46832512	0.15519568	-0.0479632	-0.0563874	-0.0431578	-0.0515816	-0.059651076
Past Sales growth rate	0.27971965	0.185727197	0.105950115	0.051077	0.46832512	1	0.19395631	0.09915826	0.02838084	0.01934918	0.03852859	0.103900642
Margin Pasado	0.121563164	0.488850548	0.862830029	0.65182788	0.15519568	0.19395631	1	0.01112399	-0.4500291	-0.4419632	-0.411717	-0.13529197
TPM Fed Hoy	0.545462427	-0.16342496	-0.283568749	-0.4473412	-0.0479632	0.09915826	0.01112399	1	0.63332674	0.55621479	0.55571464	0.640150573
Deuda Total/Total Assets	0.455497952	-0.223100023	-0.766380359	-0.8650052	-0.0563874	0.02838084	-0.4500291	0.63332674	1	0.98393954	0.9843024	0.885814678
Deuda Neta/EBITDA	0.403195715	-0.184004793	-0.759744753	-0.8640854	-0.0431578	0.01934918	-0.4419632	0.55621479	0.98393954	1	0.98656543	0.871890874
Deuda Total/Total Equity	0.41663056	-0.198823282	-0.72452356	-0.8256587	-0.0515816	0.03852859	-0.411717	0.55571464	0.9843024	0.98656543	1	0.902439153
Deuda Neta per Share	0.519258732	-0.146388705	-0.432468213	-0.5491631	-0.0596511	0.10390064	-0.135292	0.64015057	0.88581468	0.87189087	0.90243915	1

Figura 71: Matriz de Correlación.

El anexo 2 corresponde a la tabla de correlaciones de las variables macroeconómicas seleccionadas en primera instancia. Es posible apreciar que la alta correlación entre las variables de deuda y, también, la alta correlación entre SP y EP.

## 11.3. Anexo 3

Index	PMI Manu	CPI YOY Index	Earnings Pasados	Earnings growth	st Sales growth rate	Margin Pasado	TPM Fed Hoy	Deuda Neta per Share
0	54.6826	1.70522	96.4139	0.00933976	0.00305464	0.0920276	0.462696	333.47
1	46.7732	1.99756	55.7722	-0.00875663	-0.00189309	0.0736481	3.3939	473.829
2	53.7962	2.85769	55.9233	0.0101523	0.00629492	0.0832308	2.24885	638.654
3	50.78	3.288	77.0944	-0.00171616	0.00652467	0.0907007	4.2012	793.88

Figura 72: Centroides KM.

El anexo 3 corresponde a los centroides calculados por el algoritmo KM. La variable index se relaciona a cada una de las fases, recordando que: 0 – Expansión, 1 – Slowdown, 2 – Recuperación y 3 – Contracción.

## 11.4. Anexo 4



Figura 73: Evolución Probabilidades KM.



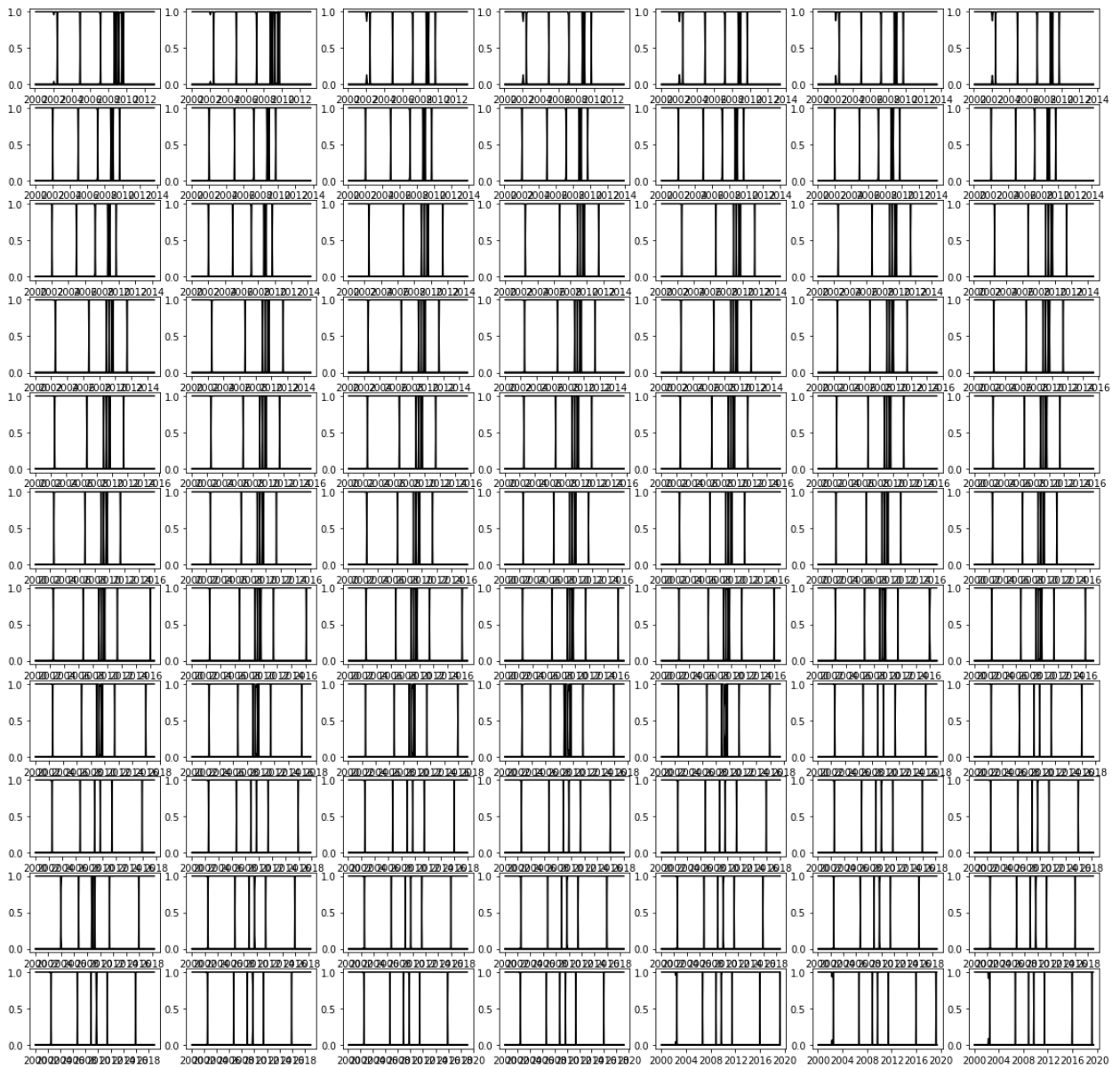
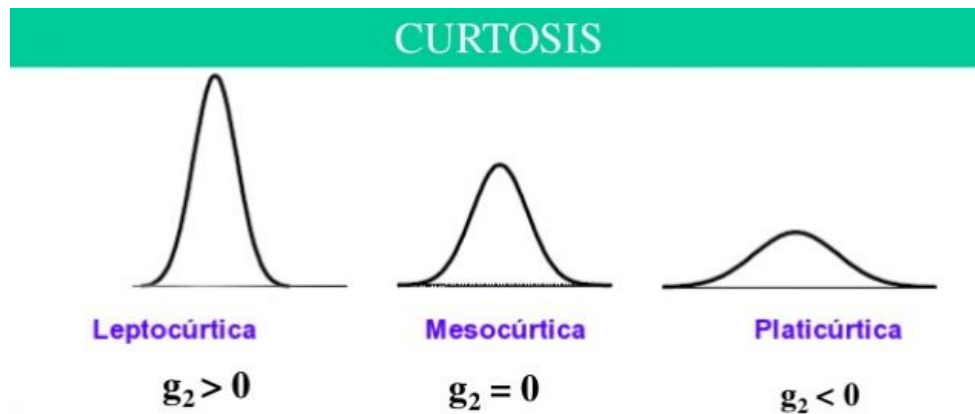


Figura 74: Evolución Probabilidades GMM.

El anexo 4 corresponde a la evolución de las probabilidades de KM y GMM para cada entrada a medida que se va agregando nueva data. Es relevante destacar que la evolución de la probabilidad, en especial en KM, no presenta grandes variaciones a medida que nueva data va ingresando, implicando que, en general, el algoritmo diferencia de buena manera las distintas fases.

## 11.5. Anexo 5



### COEFICIENTE DE CURTOSIS

Mide el grado de apuntamiento o achatamiento de la distribución de frecuencia, respecto a la curva de distribución normal que tiene coeficiente igual a 0.

$$g_2 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{s_x^4} - 3$$

Figura 75: Curtosis<sup>70</sup>.

El anexo 5 muestra lo que es la curtosis, medida que caracteriza la forma de una distribución en términos de altura.

<sup>70</sup> 100cia. Qué es la Curtosis en Estadística. Disponible en: <https://100cia.site/index.php/matematicas/item/3481-que-es-la-curtosis-en-estadistica> (consulta: 10 de enero de 2020)

## 11.6. Anexo 6

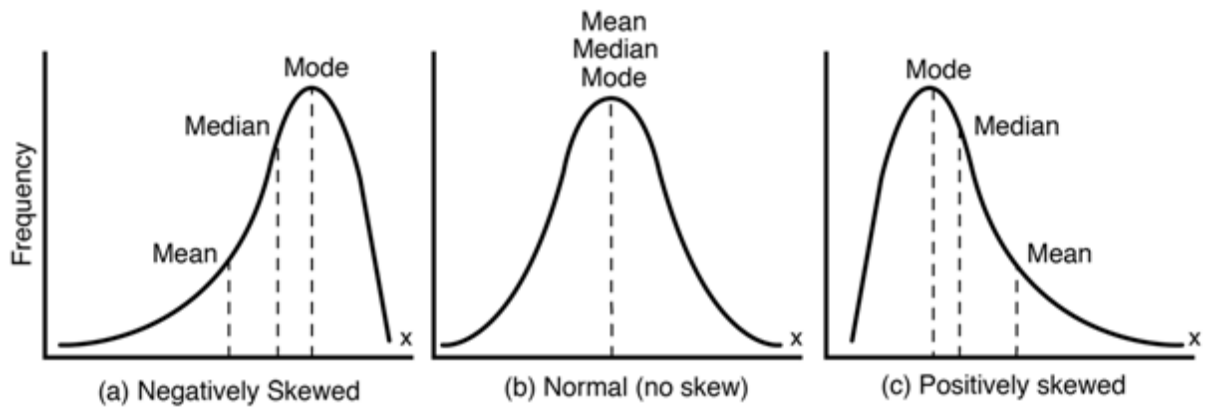


Figura 76: Skewness<sup>71</sup>.

El anexo 6 muestra lo que es la skewness, medida que caracteriza la forma de una distribución en términos de simetría.

<sup>71</sup> Researchgate. Negative/Positive Skewness. Disponible en: [https://www.researchgate.net/figure/a-Negative-skewness-b-Normal-curve-c-Positive-skewness-Durkhure-and-Lodwal-2014\\_fig5\\_294890337](https://www.researchgate.net/figure/a-Negative-skewness-b-Normal-curve-c-Positive-skewness-Durkhure-and-Lodwal-2014_fig5_294890337) (consulta: 10 de enero de 2020).