

**Tipo de documento:** Tesis de Maestría



*Departamento de Economía. Maestría en Econometría*

# Optimización de Portfolio con Preferencias de Economía del Comportamiento

**Autoría:** Shehadi Candela, Agustin Alejandro

**Año:** 2024

## ¿Cómo citar este trabajo?

Shehadi, A. (2024) "*Optimización de Portfolio con Preferencias de Economía del Comportamiento*". [Tesis de Maestría. Universidad Torcuato Di Tella]. Repositorio Digital Universidad Torcuato Di Tella <https://repositorio.utdt.edu/handle/20.500.13098/13208>

El presente documento se encuentra alojado en el Repositorio Digital de la Universidad Torcuato Di Tella bajo una licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Argentina (CC BY-NC-SA 4.0 AR)  
Dirección: <https://repositorio.utdt.edu>

# Optimización de Portfolio con Preferencias de Economía del Comportamiento

Tesis de Maestría en Econometría

**Alumno:** Agustin Alejandro Shehadi Candela

[Legajo 22W2930]

**Tutor:** Gabriel Martos Venturini, Ph.D

3 de mayo de 2024



# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Revisión de literatura</b>	<b>3</b>
2.1. Funciones de utilidad en el contexto de selección de portfolios de inversión . . . . .	4
2.1.1. Principio de preferencia por el retorno . . . . .	4
2.1.2. Principio de valoración del riesgo . . . . .	5
2.1.3. Principio de aversión a la pérdida . . . . .	6
2.1.4. Principio de valoración intertemporal . . . . .	6
<b>3. Metodología</b>	<b>7</b>
3.1. Intuición Económico-Financiera . . . . .	7
3.2. Técnicas de optimización basados en gradientes . . . . .	8
3.2.1. Background de los métodos de gradiente . . . . .	8
3.2.2. Algoritmo de descenso del gradiente . . . . .	10
3.2.3. Subgradientes . . . . .	10
<b>4. Definición de tickers y parametrización de utilidad</b>	<b>12</b>
4.1. Ejercicio de estática comparativa . . . . .	14
4.2. Análisis comparativo con métodos tradicionales . . . . .	15
4.2.1. Calibración de parámetros . . . . .	15
4.2.2. Definición de procedimiento para benchmark . . . . .	17
<b>5. Resultados</b>	<b>18</b>
<b>6. Discusión</b>	<b>20</b>
<b>7. Apéndice</b>	<b>22</b>
7.1. Principios básicos sobre funciones de utilidad . . . . .	22

## Resumen

Este trabajo presenta una alternativa innovadora en la optimización de portfolios de inversión al fusionar conceptos de finanzas bursátiles, teoría microeconómica clásica y economía del comportamiento. Nos basamos en la teoría moderna de optimización de portfolio para proponer un método que no solo busca maximizar el retorno global de la cartera y minimizar el riesgo, sino que también integra supuestos heterodoxos sobre el comportamiento de los agentes económicos. Nuestro enfoque modela la teoría del consumidor con un agente representativo que busca maximizar su utilidad, considerando tanto el retorno de su cartera como el riesgo al que está expuesto. Sin embargo, este agente también se ve restringido por hábitos, costumbres y lo que en la literatura de economía del comportamiento se conoce como *aversión a la pérdida*. Desarrollamos un algoritmo de optimización basado en subgradientes para abordar esta complejidad y encontramos una solución computacionalmente factible. Para validar nuestra propuesta, comparamos nuestros resultados con un método estándar en la práctica, el método de maximización del ratio de Sharpe, en cuatro industrias específicas: tecnológicas, bancos, gas y petróleo, y farmacéuticas. Evaluamos la alocaión óptima de activos resultante de ambos enfoques, contrastando el retorno y la volatilidad media de los portfolios obtenidos. Nuestro estudio demuestra la viabilidad y eficacia de nuestra propuesta en condiciones controladas, ofreciendo una nueva perspectiva en la optimización de portfolios de inversión.

## 1. Introducción

El propósito de este trabajo es presentar una alternativa a la literatura existente de optimización de portfolios de inversión, conectando desarrollos recientes en tres disciplinas: (i) finanzas bursátiles; (ii) teoría microeconómica clásica; y (iii) economía del comportamiento.

La teoría moderna de optimización de portfolios, sobre la que se basa este trabajo, plantea un método de alocaión de instrumentos financieros riesgosos de modo que se maximice el retorno global de la cartera, se minimice el riesgo o, en el mejor de los casos, ambas a la vez. En relación a lo anterior, este trabajo propone desarrollos novedosos para modelar la alocaión óptima de instrumentos financieros en cartera (de aquí en más, optimización de portfolio) que refleje supuestos sobre el comportamiento de los agentes económicos lo más realistas posibles. Con estos supuestos planteamos un agente representativo cuyo fin último será la maximización de su utilidad (que depende del retorno de su cartera y del riesgo al que está expuesto), pero que al mismo tiempo se ve restringido por hábitos y costumbres, y por principios de aversión a la pérdida.

Como el lector habrá advertido, la modelización de esta teoría del inversor no es sencilla, mucho menos desarrollar algoritmos que pongan en marcha estos procesos de optimización heterodoxos sobre distintas carteras de activos. Este es también un desafío que este trabajo se plantea enfrentar. Presentamos una propuesta de algoritmo de optimización basado en subgradientes, para la minimización de funciones de pérdida no estrictamente diferenciables en todo su dominio; y satisfactoriamente, llegamos a una solución computacionalmente eficiente para el problema que nos planteamos.

Finalmente, nada de este esfuerzo tendría sentido si no ponemos a prueba esta teoría contrastándola con métodos tradicionales de optimización de portfolio en condiciones controladas. Es por ello que tomamos como benchmark un procedimiento estándar en la práctica (el método de maximización del ratio de Sharpe), y comparamos contra los resultados de nuestra propuesta en cuatro industrias diferentes: (i) Tecnológicas; (ii) Bancos; (iii) Gas y Petróleo (a.k.a., Oil & Gas); y (iv) Farmacéuticas. Puntualmente, construimos cuatro portfolios de cinco activos cada uno (un portfolio por industria), y evaluamos la alocaión óptima resultante de nuestra propuesta contra la alocaión óptima resultante del benchmark. Naturalmente, la comparación última se da al contrastar el retorno y volatilidad media de los portfolios obtenidos por ambas metodologías.

El resto del trabajo se organiza de la siguiente manera. La sección 2 presenta una revisión de la literatura existente sobre procesos de optimización de portfolio y el lugar de las funciones de utilidad en esta materia. La sección 3 desarrolla la metodología a seguir bajo nuestra propuesta. La sección 4 plantea los instrumentos riesgoso seleccionados para cada industria, junto con la calibración pertinente a nuestra modelización. Finalmente, la sección 5 presenta los resultados obtenidos.

## 2. Revisión de literatura

Este trabajo asume que el lector está familiarizado con el concepto económico de funciones de utilidad. En caso que así no fuera, se presenta en el Apéndice un breve resumen de este concepto, para así

poder vincular directamente con el concepto de funciones de utilidad en el contexto de optimización de portfolios.

Respecto a la teoría moderna de gestión de portfolios, Markowitz (1952)[21] popularizó la conexión entre la teoría microeconómica sobre maximización de la utilidad y la exposición de los inversores al riesgo de portfolios de instrumentos financieros. La teoría de Markowitz construye sobre la base de lo que llamaremos **principio de preferencia por el retorno** (que discutiremos en un momento), como la maximización de una función de utilidad ( $U(R, \sigma)$ ) sobre un portfolios de activos riesgosos, que depende de su retorno global  $R$  y está restringido a una frontera de mínimo riesgo dado por la volatilidad del mismo ( $\sigma$ , e.g., varianza de cartera). La propuesta de Markowitz asume que los agentes económicos gozan de plena racionalidad a la hora de alocar su portfolios, y que estos son, a priori, neutrales al riesgo.

En este trabajo construimos sobre la base de Markowitz y la combinamos con tres ramas de la teoría económica que a priori parecerían inconexas: el **principio de valoración del riesgo** (comunmente presente en teoría económica de los seguros), el **principio de aversión a la pérdida** (presente en tópicos de economía del comportamiento), y el **principio de valoración intertemporal** (más presente en tópicos de crecimiento económico); para así llegar a una propuesta menos dependiente del supuesto de racionalidad plena de los agentes. En las siguientes subsecciones construiremos sobre los conceptos básicos de las funciones de utilidad en el contexto de selección de portfolios de inversión, continuando por la serie de principios sobre los que se construye la propuesta que traemos en el presente trabajo.

## 2.1. Funciones de utilidad en el contexto de selección de portfolios de inversión

En el contexto de alocación óptima de portfolios, la utilidad del inversor representativo sobre los activos en cartera no se plantea de manera individual, sino sobre conjunto de instrumentos como un todo. De esta manera, en cada período  $t \in T$ , el inversor evalúa los el retorno de su portfolios ( $R_t$ ) como una suma ponderada de los retornos individuales (para  $i = 1, 2, \dots, n$ ):

$$r_t(\mathbb{X}) = \sum_{i=1}^n w_{i,t} x_{i,t}$$

$$\sum_{i=1}^n w_{i,t} = 1$$

En el caso más simple se asume también que todas las ponderaciones de cartera son estrictamente *no-negativas* (i.e.,  $w_i \geq 0$  para  $i = 1, 2, \dots, n$ ), lo que implicaría que no se está permitido la operatoria en *short*.<sup>1</sup> Por simplicidad, a lo largo de este trabajo se adopta este supuesto.

### 2.1.1. Principio de preferencia por el retorno

Este principio se encuentra embebido en el trabajo seminal de Markowitz (1952)[21] y es el más intuitivo de los que presentaremos aquí. Este implica interpretar al retorno de un instrumento financiero (o conjunto de instrumentos), como un bien tradicional bajo la óptica de la función de utilidad. Dicho de otro modo, si los agentes valoran positivamente los retornos de sus instrumentos financieros decimos que su utilidad marginal respecto al retorno es positiva (i.e.,  $\frac{\partial U}{\partial r} = u'(r) > 0$ ), lo que muestra que a mayor retorno mayor será la utilidad de los agentes. No hay mucho más para decir al respecto, pero si se toma como base un conjunto de retornos de, por ejemplo, dos instrumentos financieros  $x_1$  y  $x_2$ , la utilidad asociada de esta cartera es estrictamente creciente a medida que aumenta el retorno de cualquiera de sus componentes. Como en teoría microeconómica clásica, la tenencia óptima de estos dos activos va a depender últimamente de la forma precisa de la función de utilidad y de la restricción presupuestaria que se tenga. Gráficamente:

---

<sup>1</sup>Coloquialmente, se refiere a una posición en *short* a aquella en la que el inversor toma 'prestado' un activo, sin tener propia tenencia del mismo, comprometiéndose a devolverlo a un precio pactado previamente. Desde una perspectiva económica, ya que el inversor estaría operando un activo sin tener posesión propia, se interpreta que la tenencia de este instrumento en cartera es negativo para el inversor en tiempo  $t$ .

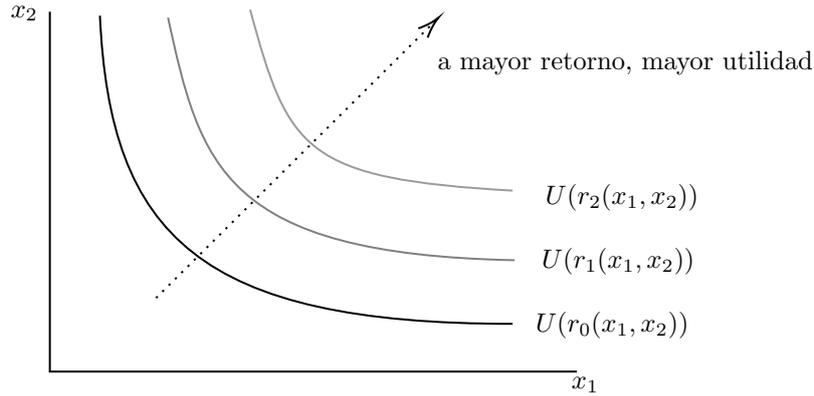


Fig.1: Ilustración de función de utilidad bajo principio de preferencia por el retorno

En donde  $r_0 < r_1 < r_2$  y podemos notar que a medida que el retorno del portfolio aumenta, también lo hace la utilidad asociada. En el caso original de Markowitz (1952),[21] se asume además que los agentes maximizan su utilidad teniendo como restricción la varianza de la cartera (si se quiere se puede interpretar este punto como un precursor al principio de valoración del riesgo en portfolios de inversión, que veremos en el apartado siguiente), para así derivar una *frontera óptima* sobre la cual los agentes económicos no pueden incrementar el retorno de su cartera sin inclinarse por un mayor apetito al riesgo.

### 2.1.2. Principio de valoración del riesgo

Tradicionalmente este concepto se limitaba a considerar al riesgo como un *mal* (en contraste a un *bien* tradicional en la teoría económica), que los agentes preferirían evitar dado que a mayor cantidades de este bien menor será su utilidad. Esto, se conoce comunmente como **aversión al riesgo** y su concepto se remonta siglos atrás por Bernoulli (1782)[2], pero no fue sino hasta Pratt (1964)[26] y Arrow (1965)[1] que, de forma independiente y casi simultánea, señalaron la importancia de la primera y segunda derivada (respectivamente) de una función de utilidad, para ver si un agente exhibe o no tendencias adversas al riesgo. El siguiente marco de Arrow-Pratt llegó a ser más tarde una base importante en la teoría financiera.

Intuitivamente, se interpreta que un agente es averso al riesgo cuando su utilidad marginal respecto al riesgo es negativa (i.e.,  $\frac{\partial U}{\partial \sigma} = u'(\sigma) < 0$ ); contrariamente, cuando esta utilidad es positiva decimos que los agentes se inclinan hacia el riesgo (a.k.a., son *amantes al riesgo*  $\frac{\partial U}{\partial \sigma} = u'(\sigma) > 0$ ).

La importancia de la segunda derivada en este esquema hace referencia a la preferencia de los individuos respecto al *outcome* esperado bajo una apuesta riesgosa. Así decimos que si la utilidad de un resultado certero ( $U(x)$ ) es mayor a la utilidad esperada de un resultado incierto ( $\mathbb{E}[U(x)|\sigma]$ ), entonces el agente es averso al riesgo.

Kimball (1990)[19] y Eeckhoudt y Schlesinger (1995)[13] se aventuraron más en este tema, y ampliaron este concepto de aversión al riesgo a las terceras derivadas de la función de utilidad, como una forma de indicar *prudencia*. Esto se relaciona entonces con la idea de que un agente preferirá un sesgo positivo en la distribución del rendimiento. La prudencia según Kimball (1990) es un concepto que establece que un agente aumentará su tasa de ahorro ante situaciones de mayor riesgo. Estos ahorros se denominan ahorros preventivos. Esto amplía el concepto de aversión al riesgo en el sentido de que si un agente solo es averso al riesgo, solo muestra aversión al riesgo, pero un agente prudente tomará medidas activas para disminuir su exposición al riesgo en tiempos de mayor volatilidad.

La actitud frente riesgo de siguiente orden se denomina *templanza*. Este se planteó por Kimball (1992)[20] y Gollier y Pratt (1996)[15] como el componente de la valoración de riesgo que desempeña un papel cuando se enfrenta un riesgo de fondo exógeno. Aunque no es tan simple como una aversión hacia la varianza, Eeckhoudt (2012)[12] afirma más precisamente que, cuando se observa la cuarta derivada como  $u^{iv}(\sigma) < 0$ , se considera efectivamente que el agente exhibe una aversión hacia la kurtosis creciente, y así sucesivamente.

A los efectos prácticos de este trabajo nos centraremos en las primeras dos derivadas para nuestra definición de función de utilidad y se deja la incorporación de eventos de orden mayor de la varianza para futuras investigaciones.

### 2.1.3. Principio de aversión a la pérdida

Kahneman y Tversky (1979)[17] realizaron investigaciones adicionales sobre cómo actúan los individuos frente a apuestas simples. En su artículo, introducen el concepto conocido como *Prospect Theory* (o Teoría de la Perspectiva en español), que gira en torno a la Aversión a las Pérdidas. La Aversión a las Pérdidas sugiere que las pérdidas tienen un mayor impacto psicológico en los individuos que las ganancias de igual magnitud. Como resultado, formalmente se interpreta que la función de utilidad propuesta en la Teoría de la Perspectiva es por tramos, con un quiebre distintivo en un punto de referencia. Para los rendimientos de activos, este quiebre se coloca típicamente alrededor de cero, ya que los cambios en la riqueza pueden ser positivos o negativos. La función de utilidad, denotada como  $u(x)$  refleja este concepto, donde  $x$  representa el cambio en la riqueza. Así  $u(x)$  se puede construir como

$$u(r) = \begin{cases} u^+(r) & \text{si } x \geq 0 \\ \lambda u^-(r) & \text{si } x < 0 \end{cases}$$

Aquí  $f(x)$  es una función estrictamente creciente y continuamente diferenciable, y  $g(x)$  comparte propiedades similares pero tiene una pendiente más pronunciada y un multiplicador  $\lambda \geq 1$ . Esta formulación resulta en una función de utilidad curvada con un quiebre ubicado en el origen, específicamente diseñada para los rendimientos de activos financieros. Gráficamente:

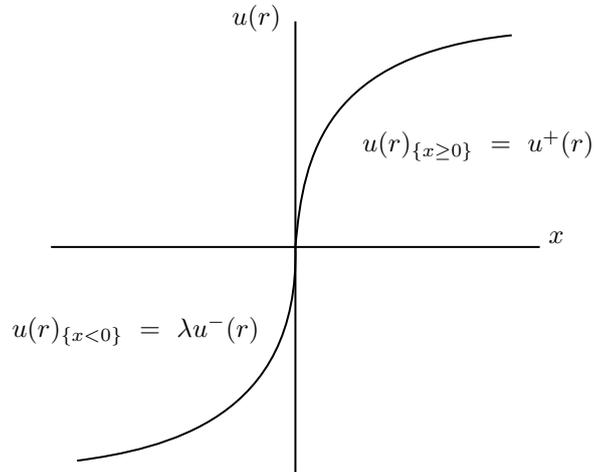


Fig.2: Ilustración de función de utilidad bajo principio de aversión a la pérdida

En el caso ilustrado se muestra un ejemplo con  $u^-(r) = -u^+(r)$  y  $\lambda = 1$ . En otras palabras, para fines ilustrativos, graficamos una función de utilidad simétrica al origen. Sin embargo, la teoría no impide que  $u^-(r)$  y  $\lambda$  tengan alguna restricción más que  $u^- : \mathbb{R} \rightarrow \mathbb{R}$  y que  $\lambda \in \mathbb{R}$ , lo que implicaría que  $u(r)$  pueda tener virtualmente cualquier pendiente y cualquier convexidad cuando  $x < 0$ .

### 2.1.4. Principio de valoración intertemporal

Finalmente, y siguiendo el desarrollo de Clauss (2013)[24], este principio es lo que subyace al concepto de valoración de *hábitos* de los agentes económicos.<sup>2</sup> La idea clave de la formación de hábitos es asumir que la utilidad asociada con la elección del consumo en una fecha determinada depende de elecciones de consumo pasadas. En un entorno de tiempo discreto, la utilidad de un proceso de consumo determinado  $c$  estaría determinado por el valor presente del flujo de consumos pasados,  $\mathbb{E}[\sum_{t=1}^T e^{\delta t} u(c_t, h_t)]$ , en donde  $h_t$

<sup>2</sup>La intuición tradicional del principio de valoración intertemporal se asocia usualmente con tópicos de crecimiento económico. En estos tipos de modelo se asume que el agente representativo tiene preferencias tanto por el consumo corriente (presente)  $c_t$  como por el consumo futuro  $c_{t+1}$ . Lo usual en este tipo de modelos es que la preferencia por el consumo corriente sea mayor que el consumo futuro; de forma tal que, parado en tiempo  $t$ , el agente representativo valora el consumo futuro mediante una *tasa de descuento* para traer este consumo futuro al presente. En el caso de los hábitos sucede exactamente lo mismo. Cuando analizamos los patrones de consumo de los agentes representativos, asumimos que el consumo presente es siempre preferible a cualquier consumo *diferido* en el tiempo. Es por esto que, al igual que cuando el agente valora el consumo futuro en menor medida comparado al consumo presente, cuando el agente valore el consumo pasado (lo que consumió con anterioridad), este *recuerdo de consumo* será valorado en menor medida que el consumo que se efectúe en el presente.

es una medida del nivel de vida o del nivel habitual de consumo. Por ejemplo,  $h_t$  puede ser un promedio ponderado de las tasas de consumo pasadas, como

$$h_t = h_0 e^{-\beta t} + \gamma \sum_{s=1}^{t-1} e^{-\beta(t-s)} c_s$$

donde  $h_0, \gamma, \beta$  son constantes no negativas. Se supone también que  $u$  disminuye en  $h$ , de modo que un alto consumo pasado genera un deseo de un alto consumo actual, es decir, las preferencias muestran complementariedad intertemporal.

La evidencia empírica sobre preferencias en la formación de hábitos es mixta. Campbell y Cochrane (1999)[7] construyen un modelo de individuo representativo con hábitos no-lineales en el que el ratio de Sharpe de equilibrio varía anticíclicamente de acuerdo con las observaciones empíricas. Sin embargo, también se puede obtener una variación contracíclica en la aversión relativa al riesgo de un individuo representativo en un modelo en el que cada individuo tiene una aversión relativa al riesgo lineal, pero las aversiones relativas al riesgo son diferentes entre los individuos, como lo explica, por ejemplo, Chan. y Kogan (2002)[9].

No obstante, varios estudios han puesto en duda la consistencia de estas presuntas preferencias por la formación de hábitos en conjuntos de datos heterogéneos (decisiones individuales sobre consumo, compras o inversiones). Por mencionar algunos estudios, Ravina (2007)[27] informa un fuerte apoyo a la formación de hábitos, mientras que Dynan (2000)[11], Gomes y Michaelides (2003)[16] y Brunnermeier y Nagel (2008)[6] no encuentran evidencia de formación de hábitos a nivel individual.

Para los fines de este trabajo se tomará como válido el supuesto de preferencias por la formación de hábitos y, en todo caso, cuando se ponga a prueba empíricamente nuestro procedimiento (contra métodos tradicionales), se definirá si se debe descartar este supuesto o no.

A continuación presentamos la propuesta de optimización de portfolio basado en los principios descritos arriba, de modo de llegar a un panorama amplio de maximización de utilidad sobre cartera de inversión, con una baja dependencia en supuestos de racionalidad y completitud de información de los agentes financieros. Es esto a lo que nos referimos cuando planteamos una propuesta de optimización de portfolios con preferencias de economía del comportamiento.

### 3. Metodología

#### 3.1. Intuición Económico-Financiera

Como se introdujo anteriormente, se plantea este estudio sobre la obra nobel de Markowitz [21], en la que se determina una frontera óptima de riesgo/retornos sobre la cual no es posible constituir una combinación lineal de activos riesgosos de mayor retorno sin aumentar el riesgo de la misma.

Formalmente, siendo  $\mathbb{X} = \{X_1, X_2, \dots, X_n\}$  una colección de activos riesgosos cuyos retornos por unidad de tiempo denotamos con la letra  $x_i$  con  $i = 1, 2, \dots, n$  y  $\mathbb{W}_i = (w_1, w_2, \dots, w_n)$  un vector de pesos no negativos tal que  $\sum_{i=1}^n w_i = 1$ , definimos el retorno de la cartera como  $r(\mathbb{X}, \mathbb{W}) = \sum_{i=1}^n w_i x_i$  y el riesgo como  $\sigma(\mathbb{X}, \mathbb{W}) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n w_i w_j x_i x_j}$ .

Así, de la mano con la teoría moderna de gestión de portfolios, el portfolio óptimo se logra cuando la función de utilidad del inversor representativo ( $U(r, \sigma)$ ) se maximiza restricta a la frontera de cartera óptima de Markowitz. Actualmente la literatura se enfocó en desarrollar calibraciones de los parámetros involucrados en  $U(r, \sigma)$  o la selección de los instrumentos a considerar, pero poco se avanzó en unificar un desarrollo de nuevas formas funcionales de  $U(r, \sigma)$  que se aparten del paradigma clásico de agentes plenamente racionales. Sobre este punto se plantea desarrollar este trabajo. Se parte de una función de utilidad dinámica no estrictamente convexa de la forma:

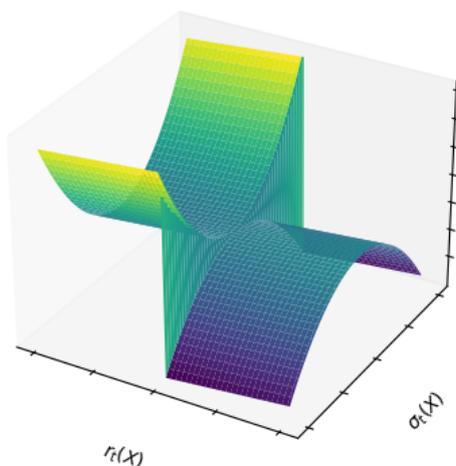
$$U(r_t, \sigma_t) = \begin{cases} r_t(\mathbb{X}, \mathbb{W}) - \frac{\alpha_1}{2} \times \sigma_{1,t}^2(\mathbb{X}, \mathbb{W}) + \beta \times U(r_{t-1}, \sigma_{t-1}) & r_t < 0 \\ r_t(\mathbb{X}, \mathbb{W}) + \frac{\alpha_2}{2} \times \sigma_{2,t}^2(\mathbb{X}, \mathbb{W}) + \beta \times U(r_{t-1}, \sigma_{t-1}) & r_t \geq 0 \end{cases} \quad (1)$$

En donde los coeficientes  $\{\alpha_1, \alpha_2\}$  representan la aversión al riesgo del inversor representativo,<sup>3</sup> y  $\beta$

<sup>3</sup>Donde  $\alpha_1$  y  $\alpha_2$  son diferentes entre sí en tanto la aversión al riesgo se magnifica diferente en función de si el individuo está ganando o perdiendo dinero. Como ejemplo, Kahneman y Tversky (1992)[18] han sugerido que la valoración de las pérdidas pueden ser dos veces más poderosas, psicológicamente, que las ganancias.

representa una suerte de preferencia intertemporal que mide el *peso* de las costumbres (a.k.a., hábitos) que tienen los agentes por sobre su utilidad pasada.<sup>4</sup> En el fondo, esta función de utilidad ilustra el principio básico de aversión a la pérdida presentado por Kahneman y Tversky [17] en el que los agentes no valoran de igual modo su aversión al riesgo cuando 'ganan dinero' respecto a cuando 'pierden dinero'. Asimismo, se plantea que el inversor representativo presente cierta valoración presente por sobre su utilidad pasada (medido por  $\beta$ ), que en el fondo implica una rigidez del inversor a la hora de actualizar los ponderadores de sus activos en cartera debido a sus hábitos y costumbres. Por completitud, se presenta debajo una ilustración de la función de utilidad planteada (para un  $t$  fijo, asumiendo  $\beta = 0$  por simplicidad), en donde podemos notar el quiebre de convexidad cuando el retorno se torna positivo. Naturalmente, con esta simplificación se deja de lado la naturaleza dinámica de  $U_t(r_t, \sigma_t)$ , pero entendemos que da noción del problema que se plantea en este trabajo

Ilustración de  $U(r_t, \sigma_t)_{\beta=0}$



Como el lector habrá advertido hasta este punto, estas no-linealidades hacen que la maximización de este tipo de utilidades no arribe a una solución analítica cerrada. Es por esto que esta tesis se plantea buscar una solución numérica para un conjunto de instrumentos y un período de tiempo dados, para así comparar los resultados obtenidos (tanto en ponderaciones óptimas como resultados de riesgo y retorno consolidado de cartera) con otros métodos ya existentes en la práctica (e.g., Sharpe Ratio maximization).

### 3.2. Técnicas de optimización basados en gradientes

Como el lector habrá podido anticipar, la maximización de una función de utilidad del tipo que se plantea en este trabajo no es una tarea posible desde el punto de vista analítico. Eso porque, dado que la utilidad depende en cada momento del tiempo del signo de  $R_t$  (a.k.a., el retorno global del portfolio en  $t$ ), su concavidad-convexidad también lo hace, y por tanto no contamos con una función estrictamente continua en todo su dominio. Dicho de otra forma, el máximo de la función no se encuentra definido para todo el dominio de sus variables constituyentes.

Es por esto que los métodos tradicionales de optimización numérica, como el descenso del gradiente *plain vanilla*, no serán de utilidad para el problema que nos enfrentamos. Así, el propósito de esta subsección es ilustrar el método del descenso del gradiente e introducir un método ajustado de este para lidiar con funciones objetivos no-estrictamente convexos o diferenciables, los *subgradiientes*.

Antes de introducir el método de los subgradiientes primero planteamos el marco teórico, las propiedades y algoritmo básico de los métodos de gradientes. Luego extenderemos esta teoría para el caso particular de los subgradiientes, que serán de crucial importancia para la solución del problema que nos planteamos en este trabajo.

#### 3.2.1. Background de los métodos de gradiente

Hoy en día hay relativo consenso en la literatura de optimización no lineal respecto a que los métodos de gradiente se deben a Louis Augustin Cauchy, en su *Compte Rendu a l'Académie des Sciences* del

<sup>4</sup>Este procedimiento sigue los lineamientos planteados por Morhaim (2023)[23] en el desarrollo matemático y propiedades óptimas de este tipo de funciones de utilidad.

18 de octubre de 1847[10]; bajo la motivación por buscar solución a cálculos pertinentes al cómputo de órbitas de cuerpos celestes. En nuestra notación contemporánea, y siguiendo los lineamientos de Boyd (2004)[5], lo que se busca es encontrar una solución al problema de optimización:<sup>5</sup>

$$\text{minimizar } f(x)$$

donde  $f : \mathcal{X} \rightarrow \mathbf{R}$ ; con  $\mathcal{X} \subseteq \mathbf{R}^n$  y  $f(x)$  es una función convexa y diferenciable. Sea  $p^* = \inf\{f(x) : x \in \mathcal{X}\}$  decimos que  $x^* \in \mathcal{X}$  es un óptimo si cumple que  $f(x^*) = p^*$ . Una condición necesaria y suficiente para que el punto  $x^*$  sea un óptimo es que

$$\nabla f(x^*) = 0$$

Para resolver este problema desde el punto de vista numérico/computacional se busca encontrar una secuencia de valores  $x^{(0)}, x^{(1)}, \dots \in \mathcal{X}$  con  $f(x^{(k)}) \rightarrow p^*$  cuando  $k \rightarrow \infty$ . Llamamos a esta la *secuencia de minimización* del problema. El algoritmo se termina cuando  $f(x^{(k)}) - p^* \geq \epsilon$  se cumple con igualdad, donde  $\epsilon > 0$  es un valor de tolerancia previamente especificado.

Los métodos de gradientes implican definir un valor de inicio  $x^{(0)} \in \mathcal{X}$  que, adicionalmente, cumpla que el conjunto  $S = \{x \in \mathcal{X} | f(x) \geq f(x^{(0)})\}$  sea un conjunto cerrado. En el caso de funciones continuas,  $\mathcal{X} = \mathbf{R}^n$  son cerrados y entonces cualquier  $x^{(0)}$  cumple la condición de pertenecer al conjunto cerrado  $S$ .

En el caso de funciones continuas y estrictamente convexas el problema es simple. En este tipo de caso, asumimos que la función es *estrictamente convexa* en  $S$ , lo que significa que existe un  $m > 0$  tal que

$$\nabla^2 f(x) \geq mI$$

Para todo  $x \in S$ . Convexidad fuerte implica consecuencias deseadas para nuestro problema. Ejemplo, para cada  $x, y \in S$  tenemos que

$$f(y) = f(x) + \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x)$$

Para algún  $z$  en el segmento lineal  $[x, y]$ . Por el supuesto de convexidad estricta último término de la derecha es, al menos,  $(m/2)\|y - x\|_2^2$ , por lo que llegamos a la siguiente desigualdad

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + (m/2)\|y - x\|_2^2$$

Para todo  $x$  e  $y$  en  $S$ . Esta última ecuación nos sirve para definir una cota sobre  $f(x) - p^*$ . Para encontrar el óptimo asumimos el gradiente respecto a  $y$  igual a cero para llegar a la expresión  $\tilde{y} = x - (1/m)\nabla f(x)$  que minimiza el término de la derecha de nuestra ecuación. Entonces tenemos que

$$\begin{aligned} f(y) &\geq f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|_2^2 \\ &\geq f(x) + \nabla f(x)^T(\tilde{y} - x) + \frac{m}{2}\|y - x\|_2^2 \\ &= f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2 \end{aligned}$$

Y dado que esto aplica para cualquier  $y \in S$ , entonces tenemos que

$$p^* \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2$$

Así, esta condición nos asegura que por la condición de continuidad y estricta convexidad de  $f(x)$  nos permita encontrar un mínimo con un cierto grado de precisión  $\epsilon$  que nosotros hayamos definido previamente:

$$\|\nabla f(x)\|_2 \leq (2m\epsilon) \Rightarrow f(x) - p^* \leq \epsilon$$

---

<sup>5</sup>El caso que se plantea aquí es el de minimización, pero los procedimientos y resultados son los mismos que en el caso de maximización. Basta con notar que el problema de minimización de  $f(x)$  es análogo al de la maximización de  $-f(x)$ .

### 3.2.2. Algoritmo de descenso del gradiente

Habiendo definido la condición de optimalidad en el apartado anterior ahora podemos definir el algoritmo típico de descenso del gradiente para problemas de optimización convexa. De nuevo, acatando el rigor detallado en Boyd (2004)[5], lo que buscamos es producir una secuencia minimizadora  $x^{(k)}$ ,  $k = 1, 2, \dots$ , donde

$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$$

Siendo el escalar  $t^{(k)} > 0$  (excepto cuando  $x^{(k)}$  es óptimo) una suerte de *tamaño de salto* entre iteración e iteración, y  $k = 1, 2, \dots$ , denota el número de iteración. Cuando en la literatura se refiere a *descenso del gradiente* nos referimos a seguir el siguiente procedimiento:

- Dado un punto de partida  $x^{(0)} \in \mathcal{X}$
- Repetimos
  1. Determinamos la dirección de descenso  $\Delta x$
  2. Elegimos un tamaño de salto  $t > 0$
  3. Actualizamos  $x := x + t\Delta x$
- Hasta que el criterio de convergencia  $|x^{(k+1)} - x^{(k)}| \leq \epsilon$  se cumpla.

Como mencionamos anteriormente, dado nuestra función objetivo (i.e., función de utilidad con preferencias de economía del comportamiento), no podemos aplicar el método del descenso del gradiente tal como se lo plantea en su concepto tradicional ya que la función de utilidad no es diferenciable. Es por eso que a continuación presentamos el concepto de subgradientes.

### 3.2.3. Subgradientes

Tomando como base un trabajo más reciente de Boyd et.al. (2022)[3], decimos que un vector  $g \in \mathbf{R}^n$  es un subgradiente de  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  en  $x \in \mathcal{X}$  si para todo  $z \in \mathcal{X}$ ,

$$f(z) \geq f(x) + g^T(z - x)$$

Si  $f$  es convexa y diferenciable, entonces el gradiente en  $x$  es un subgradiente. Pero un subgradiente puede existir incluso cuando  $f$  no sea diferenciable en  $x$ . El método numérico de optimización que se presenta en esta sección hace uso del subdiferencial; y resulta de particular relevancia en el contexto de este trabajo ya que la función de utilidad propuesta no es diferenciable en todo su dominio. Decimos que una función  $f$  es *subdiferenciable* en  $x$  si existe al menos un subgradiente de  $f$  en  $x$ . El conjunto de subgradientes de  $f$  en el punto  $x$  es llamado *subdiferencial* de  $f$  en  $x$ , y es denotada por  $\partial f(x)$ .<sup>6</sup>

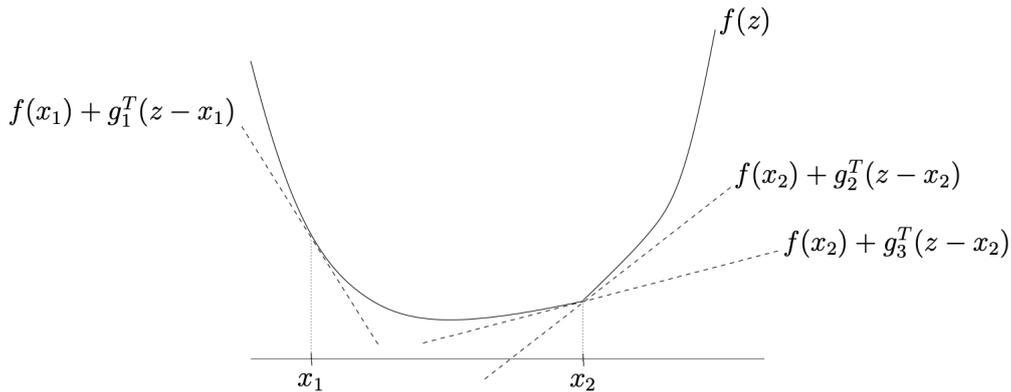
La propiedad más interesante que nos respecta sobre el subdiferencial  $\partial f(x)$  es que este es siempre un conjunto cerrado convexo, incluso si  $f$  no es convexo. Esto resulta del hecho que  $\partial f(x)$  es una intersección de un conjunto infinito de semiespacios:

$$\partial f(x) = \bigcap_{z \in \mathcal{X}} \{g \mid f(z) \geq f(x) + g^T(z - x)\}$$

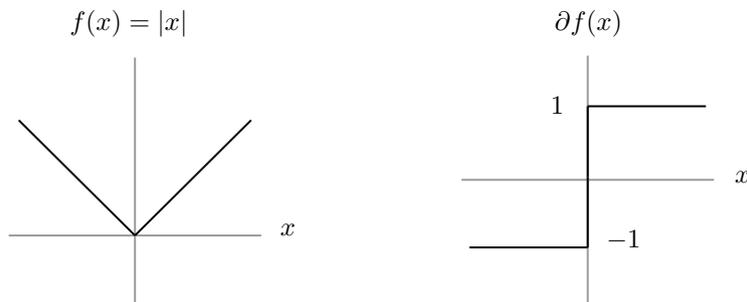
El siguiente gráfico tomado de Boyd et.al. (2022)[3] ilustra el concepto. En el punto  $x_1$ , la función convexa  $f$  es diferenciable, y  $g_1$  (que es la derivada de  $f$  en  $x_1$ ) es el único subgradiente en  $x_1$ . En cambio, en el punto  $x_2$ ,  $f$  no es diferenciable. En este punto,  $f$  tiene varios subgradientes, como por ejemplo  $g_2$  y  $g_3$ .

---

<sup>6</sup>Como ejemplo podemos proponer la función valor absoluto. Consideremos  $f(x) = |x|$ . Para  $x < 0$  el subgradiente es único:  $\partial f(x) = \{-1\}$ . Similarmente, para  $x > 0$  tenemos que el subgradiente es  $\partial f(x) = \{1\}$ . En  $x = 0$  el subdiferencial está denotado por la desigualdad  $|z| \geq gz$  para todo  $z$ , que se satisface si y solo si  $g \in [-1, 1]$ . Entonces tenemos que  $\partial f(0) = [-1, 1]$



Para un ejemplo concreto del cálculo del subgradiente se considera la función  $f(x) = |x|$ :



Habiendo introducido el concepto de subgradients, ahora podemos ilustrar un método de optimización para funciones nodiferenciables convexas. En términos de Boyd (2014)[4], este método lo denominaremos *método del subgradiente*. En el fondo este es un algoritmo bastante simple y se parece mucho al método de gradiente ordinario para funciones diferenciables, pero con varias excepciones notables:

- El método del subgradiente aplica directamente a funciones  $f$  no diferenciables
- La longitud de los pasos de iteración no se elige mediante una búsqueda lineal, como en el método de gradiente ordinario. En los casos más habituales, la longitud de los pasos se fija de antemano
- A diferencia del método de gradiente ordinario, el método de subgradiente *no es* un método de descenso; el valor de la función puede (y muy a menudo) aumentar.

El método del subgradiente que nos concierne para este trabajo toma la forma de

$$\begin{aligned} &\text{minimizar} && f_0(x) \\ &\text{sujeto a} && h_i(x) \leq 0, i = 1, 2, \dots, m \end{aligned}$$

Donde los  $h_i$  son convexas. El algoritmo toma la forma

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)}$$

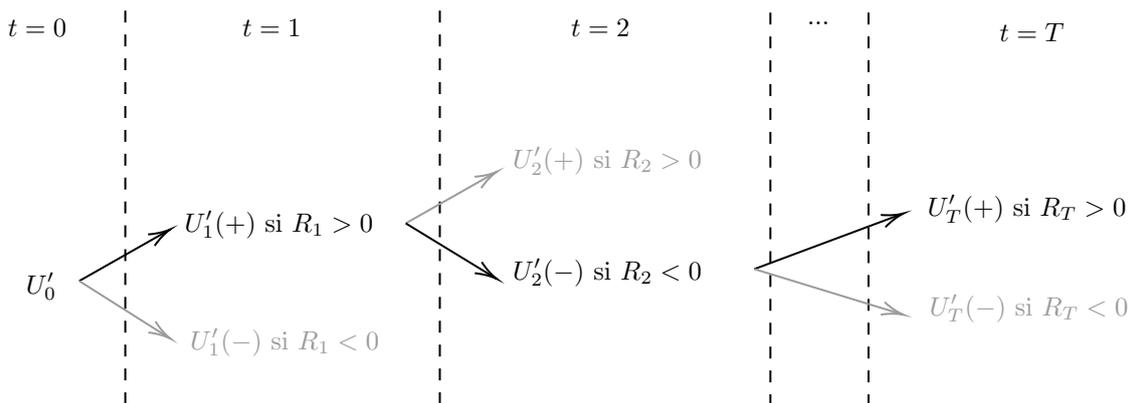
donde  $\alpha_k > 0$  es el tamaño del salto de iteración, y  $g^{(k)}$  es un subgradiente de la función objetivo en  $x^{(k)}$ . Definido como

$$g^{(k)} = \begin{cases} \partial f_0(x^{(k)}) & h_i(x^{(k)}) \leq 0, i = 1, 2, \dots, m \\ \partial f_j(x^{(k)}) & h_j(x^{(k)}) > 0 \end{cases}$$

En otras palabras: si el punto analizado es factible de diferenciar, utilizamos un subgradiente objetivo, como si el problema no tuviera restricciones; si el punto analizado no es factible, elegimos cualquier restricción violada y utilizamos un subgradiente de la función de restricción asociada. (En el último caso, podemos elegir cualquiera de las restricciones violadas, si hay más de una). En esta versión generalizada del algoritmo de subgradiente, las iteraciones pueden ser (y a menudo son) inviables. Por el contrario, las

iteraciones del método de subgradiente proyectado (y, por supuesto, el algoritmo de subgradiente básico) siempre son factibles. De esta manera, de haber un punto óptimo  $x^*$  y de cumplirse las condiciones de Slater,<sup>7</sup> decimos que hay un  $f_{\text{best}}^{(k)}$  asociado a  $x^*$  tal que  $f_{\text{best}}^{(k)} \rightarrow f^*$  en tanto  $k \rightarrow \infty$ .

En el caso puntual que nos concierne, el método del subgradiente implica tomar cada rama de la función de utilidad en tanto aplique (dependiendo si el retorno de la cartera en  $t$  sea positivo o negativo), y calcular el gradiente en dependiendo de en qué rama de la función de utilidad nos encontremos. A los fines pedagógicos, el conjunto de subgradietes que nos resultaría tendría la forma de un árbol en tanto el subgradiente en  $T$  depende de los subgradietes de todos los periodos anteriores  $t = 0, 1, \dots, T - 1$ . En línea con la intuición económico-financiera, cada vez que el subgradiente sea positivo en  $R_{t=i}$  (i.e.,  $U'_{t=i}(+)$ ) diremos que el agente es *amante al riesgo* en  $t = 1$ ; por simetría, cuando el subgradiente sea negativo en  $R_{t=j}$  (i.e.,  $U'_{t=j}(-)$ ) diremos que el agente es *averso al riesgo*.



## 4. Definición de tickers y parametrización de utilidad

Para la calibración de nuestra función de utilidad con preferencias de economía del comportamiento se toma como referencia un portfolio diversificado de cinco activos riesgosos de diferentes industrias. A saber:

### ■ Tecnológicas

- AAPL (Apple Inc.): Una empresa de tecnología multinacional que diseña, fabrica y comercializa dispositivos multimedia y de comunicaciones móviles, computadoras personales y reproductores de música digitales portátiles, y una variedad de software, servicios, periféricos, soluciones de redes y aplicaciones y contenidos digitales de terceros relacionados.
- TSLA (Tesla, Inc.): Una empresa de vehículos eléctricos y energía limpia conocida por sus automóviles eléctricos, soluciones de almacenamiento de energía y productos solares. Tesla es un destacado innovador en los sectores automotriz y de energías renovables.
- AMD (Advanced Micro Devices, Inc.): Una empresa de semiconductores especializada en el desarrollo de procesadores de computadora y tecnologías relacionadas. AMD compete con empresas como Intel y Nvidia en los mercados de CPU y GPU.
- NVDA (Nvidia Corporation): Una empresa de tecnología multinacional conocida por sus unidades de procesamiento gráfico (GPU) para juegos, visualización profesional, centros de datos y aplicaciones de inteligencia artificial (IA).
- GOOGL (Alphabet Inc.): La empresa matriz de Google, que ofrece una amplia gama de productos y servicios que incluyen búsqueda en Internet, publicidad en línea, computación en la nube y productos de hardware como teléfonos inteligentes Pixel y Google Home.

### ■ Bancos

<sup>7</sup>i.e., el problema se llama estrictamente factible en tanto se pueda definir un  $f_{\text{best}}^{(k)} = \min f_0(x^{(i)}) | x^{(i)} \text{ factible}, i = 1, 2, \dots, k$ .

- JPM (JPMorgan Chase & Co.): Una importante firma global de servicios financieros que ofrece una variedad de servicios bancarios, de inversión y de gestión de activos. JPMorgan Chase es uno de los bancos más grandes de los Estados Unidos.
- BAC (Bank of America Corporation): Un banco de inversión multinacional y empresa de servicios financieros que ofrece servicios bancarios, de inversión y de gestión de patrimonio a individuos, empresas e instituciones.
- WFC (Wells Fargo & Company): Una empresa diversificada de servicios financieros que ofrece servicios bancarios, de inversión, hipotecarios y de financiamiento al consumidor. Wells Fargo es uno de los bancos más grandes del mundo por capitalización de mercado.
- C (Citigroup Inc.): Una empresa global de servicios financieros que ofrece servicios bancarios, de inversión y de seguros a individuos, corporaciones y gobiernos en todo el mundo.
- GS (The Goldman Sachs Group, Inc.): Una destacada firma global de banca de inversión, valores y gestión de inversiones que ofrece una amplia gama de servicios financieros a corporaciones, gobiernos e individuos.

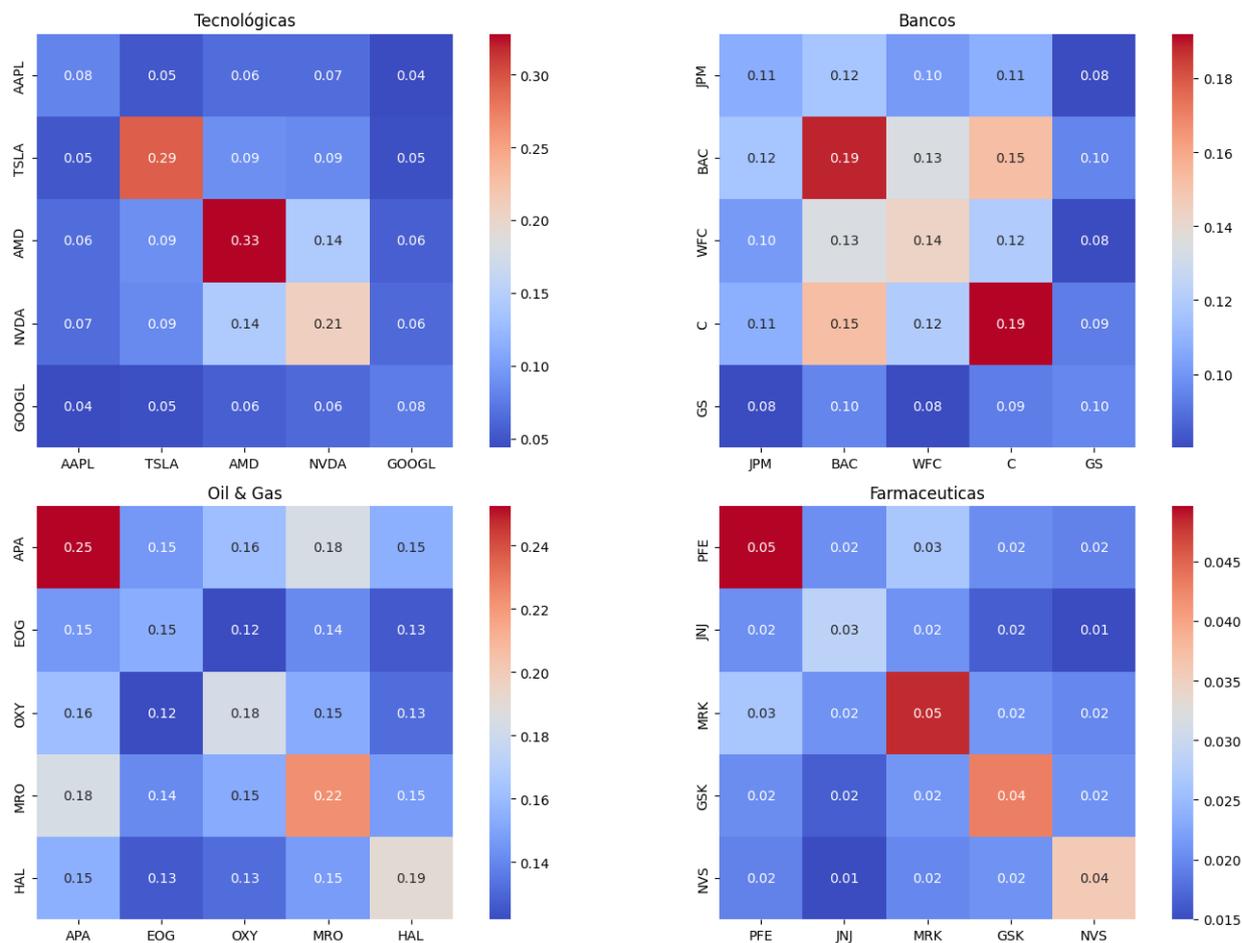
#### ■ Oil & Gas

- APA (Apache Corporation): Apache Corporation es una empresa petrolera estadounidense con sede en Houston, Texas. Se dedica a la exploración, producción y comercialización de petróleo y gas natural en varios países del mundo.
- EOG (EOG Resources, Inc.): EOG Resources es una empresa de exploración y producción de petróleo y gas con sede en Houston, Texas. Se centra en la exploración, desarrollo, producción y comercialización de recursos de petróleo y gas natural en los Estados Unidos.
- OXY (Occidental Petroleum Corporation): Occidental Petroleum Corporation es una empresa petrolera multinacional con sede en Houston, Texas. Opera en la exploración, producción y comercialización de petróleo, gas natural y productos químicos en todo el mundo.
- MRO (Marathon Oil Corporation): Marathon Oil Corporation es una empresa petrolera estadounidense con sede en Houston, Texas. Se dedica a la exploración, producción y comercialización de petróleo y gas natural en los Estados Unidos y otros países.
- HAL (Halliburton Company): Halliburton es una empresa de servicios petroleros con sede en Houston, Texas. Proporciona una amplia gama de servicios y productos para la industria del petróleo y gas, incluyendo servicios de perforación, completación y producción.

#### ■ Farmacéuticas

- PFE (Pfizer Inc.): Una corporación farmacéutica multinacional dedicada al descubrimiento, desarrollo, fabricación y comercialización de productos farmacéuticos, incluidas vacunas, medicamentos y productos de atención médica para consumidores.
- JNJ (Johnson & Johnson): Una empresa multinacional de dispositivos médicos, productos farmacéuticos y bienes de consumo conocida por sus productos de salud, que incluyen medicamentos con receta, dispositivos médicos y productos de salud para consumidores como Band-Aid y Tylenol.
- MRK (Merck & Co., Inc.): Una empresa global de atención médica dedicada al descubrimiento, desarrollo, fabricación y comercialización de una amplia gama de productos farmacéuticos y vacunas para la salud humana y animal.
- GSK (GlaxoSmithKline plc): Una empresa farmacéutica multinacional británica dedicada a la investigación, desarrollo, fabricación y comercialización de productos farmacéuticos, vacunas y productos de atención médica para consumidores.
- NVS (Novartis AG): Una empresa farmacéutica multinacional suiza dedicada a la investigación, desarrollo, fabricación y comercialización de productos farmacéuticos, medicamentos genéricos, medicamentos de venta libre y productos para el cuidado de los ojos.

Una vez definido los portfolios sobre los que vamos a trabajar, procedemos a analizar la matriz de varianzas y covarianzas de los activos elegidos para cada industria. Tomando como referencia el período temporal comprendido entre enero 1, 2009 y enero 1, 2024 (se toma este período de tiempo para cubrir 15 años de información), la matriz de varianzas y covarianzas de los retornos diarios de los componentes se ilustra en la siguiente figura:



Algo notable a destacar a primera vista es el nivel relativo de volatilidad entre las industrias. Es bastante evidente que la volatilidad media de la industria tech es la mayor de todas; luego le sigue la industria del Oil & Gas, seguido por el sector bancario y últimamente el sector farmacéutico. Otro hecho notorio es, dado los tickers seleccionados, el sector en el que las empresas más se parecen entre sí es el sector bancario (i.e., las covarianzas son relativamente homogéneas entre sí). En contraste tenemos al sector tecnológico, con firmas extremadamente volátiles como AMD y firmas relativamente estables como Google.

Habiendo definido el conjunto de tickers e industrias a analizar, procederemos en primer lugar a desarrollar un análisis de estática comparativa para estudiar la sensibilidad de nuestro procedimiento ante variaciones de los parámetros relevantes. Dejamos para sección 4.2 la calibración de los parámetros de la función de utilidad propuesta y la contrastación con otros métodos.

#### 4.1. Ejercicio de estática comparativa

Para este análisis tomaremos como objeto de estudio el portfolio de acciones farmacéuticas. Procederemos a aplicar nuestro algoritmo de optimización de funciones de utilidad de economía del comportamiento, haciendo variar tanto  $\alpha_1$  como  $\alpha_2$  y  $\beta$ , y registrando el resultante retorno y volatilidad de los portfolios óptimos para cada parametrización.<sup>8</sup>

Primeramente vamos a estudiar nuestro algoritmo tomando  $\alpha_1^{(i)} \in \{1, 2, 3, 4, 5\}$ ,  $\alpha_2 = 2$ , y  $\beta = 0,9$ . Debajo una tabla mostrando los resultados. Desafortunadamente, no se aprecia a simple vista una relación clara entre retorno y volatilidad óptimo a medida que aumentamos el  $\alpha_1$ . En otras palabras, en medida que aumentamos el parámetro de aversión al riesgo no encontramos una relación clara entre retorno y volatilidad óptimos del portfolio. Una primera hipótesis puede ser que este fenómeno se debe al conjunto específico de acciones/industria que se tomó como referencia. No obstante, mayores estudios sobre la

<sup>8</sup>En la siguiente subsección se calibrarán los parámetros acorde a la evidencia empírica. Hasta entonces, la presente subsección presenta un ejercicio de carácter más bien analítico sobre el procedimiento que se plantea en este trabajo.

sensibilidad de este tipo de funciones de utilidad respecto al parámetro de aversión al riesgo se deja para futuras investigaciones.

$\alpha_1$	Return (%)	Volatility (%)
1	17.74 %	15.65 %
2	17.67 %	15.57 %
3	16.72 %	14.89 %
4	17.61 %	15.48 %
5	17.50 %	16.87 %

Ahora llevaremos a cabo el mismo ejercicio, pero en lugar de hacer variar el  $\alpha_1$ , haremos variar  $\alpha_2$ . Específicamente, asumimos  $\alpha_1 = 2$ ,  $\alpha_2^{(i)} \in \{1, 2, 3, 4, 5\}$ , y  $\beta = 0,9$ . Debajo los resultados. En este caso sí se aprecia una relación clara entre el parámetro  $\alpha_2$  y los retornos y volatilidad óptimos. Como cabría esperarse, a medida que aumenta la aversión a la pérdida notamos que aumenta la volatilidad esperada del portfolio. Asimismo, a medida que aumenta el riesgo esperado de portfolio también aumenta el retorno esperado.

$\alpha_2$	Return (%)	Volatility (%)
1	17.45 %	15.29 %
2	17.67 %	15.57 %
3	17.74 %	15.65 %
4	17.76 %	15.68 %
5	17.76 %	15.68 %

Finalmente, estudiaremos cómo varía el retorno y volatilidad óptimos en función de variaciones de  $\beta$ . Asumimos,  $\alpha_1 = \alpha_2 = 2$  y  $\beta^{(i)} \in \{0,1, 0,25, 0,5, 0,75, 1\}$ . En este caso agregamos decimales porque las diferencias son más sutiles que en las comparaciones anteriores. En esta comparación de estática comparativa notamos que, llamativamente, a medida que aumenta la tasa de preferencia intertemporal, el efecto en retorno disminuye, pero el efecto en volatilidad no es 100 % certero. En otras palabras, a priori parecería que la volatilidad baja a medida que aumenta  $\beta$ , pero la última observación (i.e., cuando  $\beta = 1$ ) sugeriría que la volatilidad aumenta también. Mayor profundidad en estudiar estos fenómenos se deja para futuras investigaciones.

$\beta$	Return (%)	Volatility (%)
0.1	17.6726	15.5653
0.25	17.6724	15.565
0.5	17.6722	15.5649
0.75	17.6719	15.5644
1.0	17.7644	15.6861

## 4.2. Análisis comparativo con métodos tradicionales

### 4.2.1. Calibración de parámetros

Para comparar la performance de nuestro procedimiento contra otros métodos de selección de cartera tradicionales usados en la industria, tenemos que definir una parametrización para la función de utilidad acorde al sustento económico empírico.

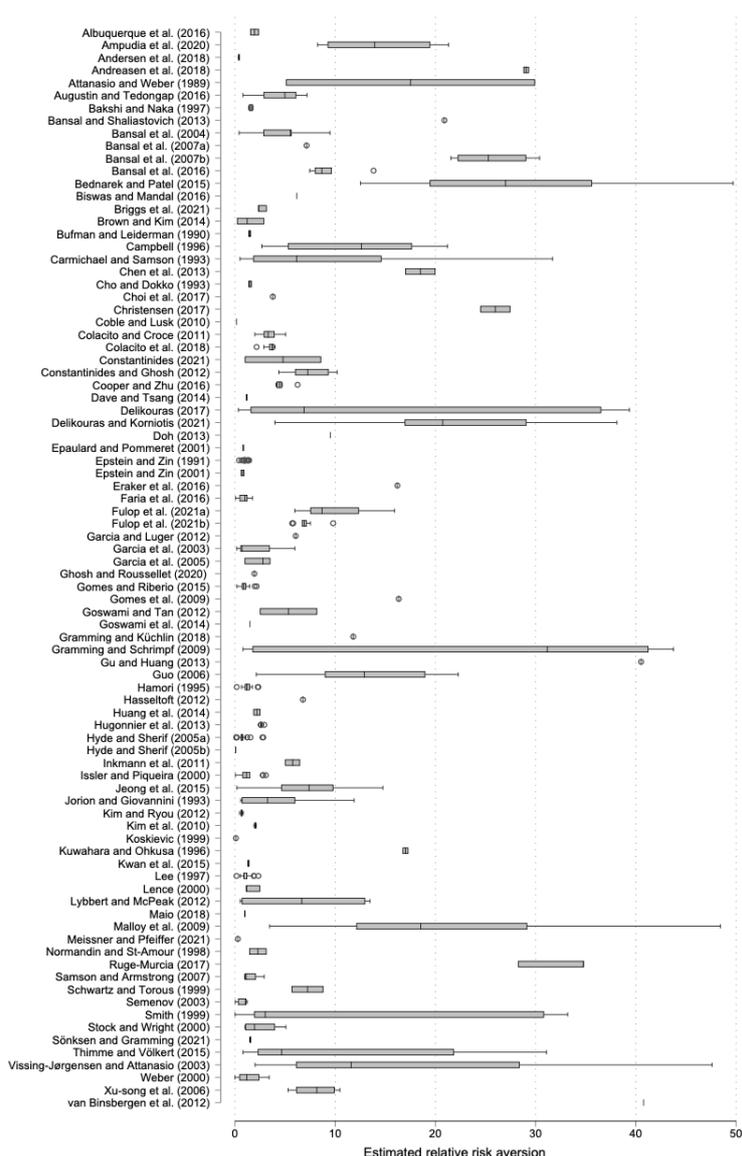
Vamos a comenzar con el parámetro de preferencia intertemporal  $\beta$  que, como se definió en la subsección de principio de valoración intertemporal, se puede interpretar como el factor de descuento en tiempo discreto. En notación de Clauss (2013)[24], y asumiendo un valor inicial  $h_0 e^{-\beta t} = 0$ , definimos nuestra calibración de  $\beta$  como  $\tilde{\beta}$  y la asumimos como la inversa de la tasa libre de riesgo ( $\tau_{\text{risk free}}$ ) a la fecha de corte del dataset (i.e., enero 1 de 2024):<sup>9</sup>

$$\tilde{\beta} = \gamma e^{-\beta(t-s)} = \frac{1}{\tau_{\text{risk free}}} = \frac{1}{1,0408}$$

En lo que refiere al parámetro de aversión al riesgo (i.e.,  $\alpha_1$  de la ecuación 1) se toma como referencia el trabajo de Elminejad (2022) [14]. Este consiste en un meta-análisis sobre un total de 1.021 estimaciones

<sup>9</sup>Se toma como referencia el retorno de los bonos soberanos de los Estados Unidos a 30 años con maturity constante. Disponible en <https://fred.stlouisfed.org/series/DGS30>

(resultante de 92 trabajos) sobre la medición de el parámetro de aversión al riesgo. La figura debajo es un extracto de la Figura 2 del paper de Elminejad (2022) e ilustra la distribución 'cruda' de los parámetros de aversión al riesgo que tienen en cuenta en su dataset:



Sobre este conjunto de parámetros, el autor identifica la presencia de un sesgo de publicación,<sup>10</sup> y heterogeneidad temporal;<sup>11</sup> para finalmente aplicar un método de *Bayesian Model Averaging* para dar con intervalos de confianza de una estimación insesgada del parámetro de aversión al riesgo. Como resultado de su meta-análisis, y tras haber ajustado su procedimiento por los sesgos pertinentes, Elminejad (2022) [14] reporta en su Tabla 6 sus resultados ajustados. Aquí debajo un extracto de dicha tabla:

<sup>10</sup>El problema es que, si bien es difícil identificar *estimaciones atípicamente grandes* (no existe a priori un umbral superior para la aversión al riesgo), los investigadores usualmente descartan (o evitan publicar) aquellas estimaciones que son negativas o estadísticamente insignificantes. Ante resultados poco intuitivos, los investigadores pueden optar por no informarlos o probar una especificación diferente con la esperanza de obtener resultados que sean consistentes con los antecedentes de la literatura. Este tipo de censura tiende a elevar virtualmente la aversión al riesgo media reportada respecto del valor real, y esto es lo que los meta-analistas llaman sesgo de publicación (Card y Krueger (1995);[8] Stanley (2001)[29]).

<sup>11</sup>Algunos de las estimaciones se lograron a través de series temporales, otras de datos de panel o de cortes transversales. Adicionalmente, de aquellos estudios que se hacen a través de series temporales algunas fueron realizadas mediante datos diarios, otros trimestrales o incluso anuales. A esta heterogeneidad el autor se refiere como *heterogeneidad temporal*

Table 6: Implied risk aversion

	Mean	95% cred. int.
Overall best practice	3.73	[-7.36, 14.82]
Economics	1.24	[-10.25, 12.73]
Finance	7.16	[-3.85, 18.17]
US	5.81	[-5.64, 17.26]
EU	1.57	[-7.07, 10.22]
Stockholder	1.49	[-6.80, 9.79]
GMM	3.79	[-6.94, 14.52]
Quarterly data	6.33	[-4.61, 17.27]

*Notes:* The table uses benchmark BMA results to compute relative risk aversion conditional on selected aspects of data, methodology, and publication (see text for details). That is, the table attempts to answer the question what the mean risk aversion would look like if the literature was free of publication bias and all studies used the same strategy as the one we prefer. The 95% credible intervals are reported in parentheses.

Para los fines del presente trabajo tomaremos como parámetro de aversión al riesgo de nuestra función de utilidad al correspondiente a la categoría *Overall best practices* de Elminejad (2022) [14]. Por tanto, asumiremos  $\alpha_1 = 3,73$ .

Para completar la parametrización de nuestra función de utilidad (i.e., Ecuación 1), solo resta definir el parámetro de *no-aversión al riesgo*,  $\alpha_2$  (a.k.a., factor de aversión a la pérdida). Para ello vamos a tomar como base el trabajo de Pachur & Kellen (2013)[25]. En este, los autores testean sistemáticamente distintas especificaciones de modelos para estimar el factor de aversión a la pérdida y, a diferencia de literatura actual que citan (y en favor del procedimiento presente en este trabajo), los autores llegan a la conclusión de que el efecto de la aversión a la pérdida existe y es relevante, pero no es necesario asumir formas funcionales diferentes cuando se trata de una apuesta riesgosa. En su nomenclatura, definen  $\lambda$  como la sensibilidad relativa a las pérdidas frente a las ganancias; y mencionan que cuando esta es mayor que 1, se indica aversión a las pérdidas. Visto de otro modo, puede interpretarse a  $\lambda$  como un escalador que, de existir, magnifica la utilidad cuando los individuos se enfrentan a apuestas riesgosas. En su trabajo, Pachur & Kellen (2013)[25] identifican un  $\lambda = 1,4$  significativamente distinto de 1; lo que, de adoptar este parámetro para nuestras estimaciones, implicaría que

$$\alpha_2 = \lambda \alpha_1 = 1,4 \times 3,73 = 5,222$$

A modo de resumen, la tabla debajo presenta la calibración de parámetros que tomaremos para la estimación de nuestra función de utilidad:

$\beta$	$\alpha_1$	$\alpha_2$
1	3,73	5,22
1,0408		

#### 4.2.2. Definición de procedimiento para benchmark

Para establecer un benchmark antes de evaluar nuestro proceso de optimización basado en función de utilidad, primero vamos a calcular un portofolio óptimo siguiendo prácticas usuales de la industria. Como benchmark vamos a tomar el procedimiento de maximización de ratio de Sharpe ( $S$ ). Este es un índice muy utilizado en la práctica de gestión de riesgos bursátiles y se define como

$$S = \frac{r(\mathbb{X}, \mathbb{W}) - r_f}{\sigma^2(\mathbb{X}, \mathbb{W})}$$

Donde  $r(\mathbb{X}, \mathbb{W})$  es el retorno del portofolio de activos riesgosos  $\mathbb{X}$ ,  $r_f$  es el retorno libre de riesgos (usualmente el rendimiento de los bonos soberanos americanos a 30 años), y  $\sigma^2(\mathbb{X}, \mathbb{W})$  es la varianza de portofolio. Recordando que el retorno y la varianza de portofolio son sumas ponderadas de los retornos/varianzas de los activos constituyentes, el objetivo de este procedimiento es encontrar el conjunto de ponderaciones óptimas de cartera  $\mathbb{W}^*$  definido como

$$\mathbb{W}^* = \operatorname{argmax} \frac{r(\mathbb{X}, \mathbb{W}) - r_f}{\sigma^2(\mathbb{X}, \mathbb{W})}$$

$$\begin{aligned} \text{s.t. } r(\mathbb{X}, \mathbb{W}) &= \sum_{i=1}^n w_i x_i \\ \sigma^2(\mathbb{X}, \mathbb{W}) &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j x_i x_j \\ 1 &= \sum_{i=1}^n w_i \end{aligned}$$

Para este tipo de procedimientos existen algoritmos finamente optimizados para dar con el resultado en el menor tiempo posible.<sup>12</sup> Así, tomaremos los cuatro portfolios descritos en la sección anterior para llegar a las ponderaciones óptimas para cada industria según el procedimiento de maximización del ratio de Sharpe. Con estas ponderaciones óptimas calcularemos el retorno y volatilidad promedio que arrojarían estas carteras. Finalmente calcularemos las ponderaciones óptimas (y el consecuente par retorno-volatilidad) resultantes de correr el procedimiento descrito en este trabajo (i.e., maximización de la *behavioral utility*) para determinar industria por industria si nuestro procedimiento mejora (o no) los resultados obtenidos por métodos convencionales de la industria.

## 5. Resultados

Presentamos debajo las ponderaciones óptimas de benchmark para las cuatro industrias, según el método de maximización del ratio de Sharpe.

Tech		Banking	
Ticker	Weight (%)	Ticker	Weight (%)
AAPL	35.3 %	JPM	59.6 %
TSLA	20.8 %	BAC	11.3 %
AMD	3.2 %	WFC	0.0 %
NVDA	32.4 %	C	0.0 %
GOOGL	8.3 %	GS	29.2 %

Oil & Gas		Pharma	
Ticker	Weight (%)	Ticker	Weight (%)
APA	0.0 %	PFE	11.6 %
EOG	95.9 %	JNJ	25.0 %
OXY	0.0 %	MRK	38.4 %
MRO	0.0 %	GSK	0.0 %
HAL	4.1 %	NVS	25.0 %

Así, si utilizamos estas ponderaciones óptimas (según el procedimiento de maximización del ratio de Sharpe) para calcular el retorno y volatilidad media para cada industria, llegamos al siguiente resultado:

Resumen: Retorno esperado y volatilidad - Benchmark		
Industry	Expected annual return	Annual volatility
Tech	38.2 %	30.5 %
Banking	15.2 %	31.4 %
Oil & Gas	10.5 %	38.9 %
Pharma	11.7 %	16.2 %

En este punto ya tenemos todos los ingredientes para poner a prueba el procedimiento que se propone en este trabajo. En otras palabras, calcularemos el portfolio óptimo según la maximización de utilidad de comportamiento (ver Ecuación (1)), para las cuatro industrias propuestas en la sección anterior, y compararemos la distribución óptima de instrumentos y el respectivo retorno-volatilidad promedio contra el benchmark.

En las tablas debajo se busca precisamente esto, presentamos las ponderaciones óptimas según nuestra propuesta par-a-par con las resultantes del procedimiento de maximización del ratio de Sharpe.

<sup>12</sup>En el caso puntual de este trabajo se usó el módulo **EfficientFrontier** de la librería **pypfopt** de python.

Resultados: Ponderaciones Tech - Max. Utilidad de comportamiento

Ticker Name	Max. Sharpe Ratio	Max. Behavioral Utility
AAPL	35.3 %	15.8 %
TSLA	20.8 %	22.1 %
AMD	3.2 %	24.8 %
NVDA	32.4 %	22.1 %
GOOGL	8.3 %	15.2 %

Resultados: Ponderaciones Bancos - Max. Utilidad de comportamiento

Ticker Name	Max. Sharpe Ratio	Max. Behavioral Utility
JPM	59.6 %	18.9 %
BAC	11.3 %	21.9 %
WFC	0.0 %	20.1 %
C	0.0 %	21.6 %
GS	29.2 %	17.5 %

Resultados: Ponderaciones Oil & Gas - Max. Utilidad de comportamiento

Ticker Name	Max. Sharpe Ratio	Max. Behavioral Utility
APA	0.0 %	21.4 %
EOG	95.9 %	18.7 %
OXY	0.0 %	19.5 %
MRO	0.0 %	20.8 %
HAL	4.1 %	19.5 %

Resultados: Ponderaciones Farmacéuticas - Max. Utilidad de comportamiento

Ticker Name	Max. Sharpe Ratio	Max. Behavioral Utility
PFE	11.6 %	21.3 %
JNJ	25.0 %	18.2 %
MRK	38.4 %	21.3 %
GSK	0.0 %	20.0 %
NVS	25.0 %	19.1 %

Una conclusión que salta a la vista es que nuestro procedimiento parecería sugerir carteras más diversificadas. Esto en tanto las ponderaciones dentro de cada cartera son más homogéneas, comparado al procedimiento de maximización del ratio de Sharpe. Formalmente, como medida de 'grado de diversificación' tomamos la métrica de *entropía*. En teoría de la información,<sup>13</sup> la entropía de una variable aleatoria se define como el nivel promedio de dispersión o 'sorpresa' inherente a las posibles realizaciones de la variable. Dada una variable aleatoria discreta  $z$  con soporte en el conjunto  $\text{sup}(z)$  y que se distribuye acorde a  $p(z)$ , la entropía se calcula con

$$H(z) := - \sum_{z \in \text{sup}(z)} p(z) \log p(z)$$

En el caso que nos concierne, mayor entropía de un portfolio denotaría un mayor grado de diversificación (i.e., las ponderaciones de los constituyentes se distribuyen de manera más *homogénea* dentro del portfolio). La tabla debajo muestra el resultado de la entropía de los distintos portfolios óptimos según el benchmark (i.e., maximización del ratio de Sharpe) comparado con nuestro procedimiento propuesto. Como se puede observar, las carteras seleccionadas con los métodos propuestos en este trabajo han resultado más diversificadas en los ejemplos analizados en esta sección.

Industry	Sharpe ratio	Behavioral utility
Tech	1.3759	1.5907
Banking	0.9139	1.6059
Oil & Gas	0.1726	1.6082
Pharma	1.3108	1.6075

<sup>13</sup>El concepto de entropía fue introducido por Claude Shannon en su artículo de 1948 *Mathematical Theory of Communication*[28], y también se conoce como entropía de Shannon.

Finalmente, bajo nuestra propuesta arribamos a siguiente conjunto de retornos y volatilidades. Es notorio, cuando se compara con los resultados del benchmark, que nuestro procedimiento parecería sugerir carteras más diversificadas y con un retorno promedio mucho más elevado (con un sacrificio en riesgo relativamente bajo). Esto independientemente de las industrias que tomamos como referencia. Lo que es llamativo, y se deja abierto para futuras investigaciones, es que para la industria farmacéutica nuestro procedimiento no solo arroja un mayor retorno promedio comparado contra el benchmark, sino también ofrece una volatilidad promedio menor.

Expected performance of optimal portfolio through Behavioral Utility maximization

Industry	Expected annual return	Annual volatility
Tech	78.0 %	33.0 %
Banking	29.7 %	34.4 %
Oil & Gas	23.0 %	39.9 %
Pharma	17.8 %	15.7 %

## 6. Discusión

El presente trabajo se propuso modelar la optimización de portfolios de inversión incorporando supuestos heterodoxos sobre el comportamiento de los agentes económicos. En lo puntual se desarrolló un modelo que asume que los inversores tienen preferencias asimétricas respecto a la valoración de riesgos (siendo estos aversos al riesgo para las pérdidas de capital y amantes al riesgo para las ganancias), al mismo tiempo que tienen valoración por la formación de hábitos y costumbres (a.k.a., preferencias intertemporales). Desde el punto de vista analítico, esto plantea un desafío al no poder definir la solución cerrada desde el punto de vista matemático. Así, para dar con una solución, se debe acudir a métodos numéricos de optimización por subgradientes para resolver el problema.

Una vez superadas las barreras teóricas y computacionales se procedió a evaluar la performance de nuestro procedimiento contra otro utilizado comúnmente en la práctica profesional. Se optó así por comparar contra el procedimiento de maximización del ratio de Sharpe. Consideramos cuatro industrias (tecnológicas, bancos, oil & gas y farmacéuticas), con cinco activos cada uno, para luego comparar la combinación óptima resultante de los dos procedimientos con su respectiva combinación de riesgo/retorno. En principio, el resultado al que arribamos sugeriría que nuestro procedimiento es superior al benchmark. Esto es porque, dadas las industrias consideradas y el conjunto de activos que las constituyen, nuestro procedimiento daría carteras más diversificadas (i.e., con mayor entropía), con mayor retorno promedio y con un sacrificio relativamente bajo en términos de volatilidad media.

No obstante, hay algunos puntos que se dejan planteados para futuras investigaciones. En primer lugar, se abre la posibilidad de estudiar la estabilidad de nuestros resultados tomando carteras con mayor número de activos constituyentes o incluso de un mayor número de industrias. Segundo, queda pendiente un estudio más profundo y exhaustivo sobre la función de utilidad asumida en este trabajo. En este trabajo se tocó de manera superficial un análisis de estática comparativa sobre los parámetros involucrados en la función de utilidad, pero sin resultados concluyentes. Finalmente, si bien no se dedicó un espacio puntual en este trabajo para desarrollarlo, queda pendiente el estudio fino del costo computacional de llevar adelante un procedimiento como el que aquí se plantea. Particularmente, cuando se evaluó el procedimiento de optimización de benchmark, el tiempo medio de cómputo fue de 0.02 segundos. En contraste, nuestro procedimiento demostró ser mucho más demandante desde el punto de vista computacional, elevando el tiempo medio de cómputo a 18.5 segundos. Se deja abierta la posibilidad de plantear algoritmos más optimizados en lo que al tiempo de cómputo respecta.

## Referencias

- [1] Kenneth Arrow. *The theory of risk aversion*. Yrjo Jahnssonin Saatio, Helsinki, 1965. Reprinted in: *Essays in the Theory of Risk Bearing*.
- [2] Daniele Bernoulli. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperiales Petropolitane*, pages 175–192, 1782.
- [3] S. Boyd, J. Duchi, M. Pilanci, and L. Vandenberghe. Subgradients. Notes for EE364b, Stanford University, Spring 2021-22, April 2022. Accessed on April 13, 2022.
- [4] Stephen Boyd. Subgradient methods. Notes for EE364b, Stanford University, Spring 2013–14, 2013.
- [5] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [6] Markus K. Brunnermeier and Lasse H. Pedersen. Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6):2201–2238, 2009.
- [7] John Y. Campbell and John H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251, 1999.
- [8] D. Card and A. B. Krueger. Time-series minimum-wage studies: A meta-analysis. *American Economic Review*, 85(2):238–243, 1995.
- [9] Yuk Leung Chan and Leonid Kogan. Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy*, 110(6):1255–1285, 2002.
- [10] Jean Dhombres and Christian Gilain. Bibliographie concernant cauchy (1974 à aujourd’hui). *Revue d’histoire Des Sciences*, 45(1):129–133, 1992. Accessed 30 Mar. 2024.
- [11] Karen E. Dynan. Habit formation in consumer preferences: Evidence from panel data. *American Economic Review*, 90(3):391–406, 2000.
- [12] Louis Eeckhoudt. Beyond risk aversion: Why, how and what’s next? *The Geneva Risk and Insurance Review*, 37:141–155, 2012.
- [13] Louis Eeckhoudt and Harris Schlesinger. Higher-order risk attitudes. In *Handbook of Insurance*, pages 41–57. Springer New York, 1995.
- [14] Ali Elminejad, Tomas Havranek, and Zuzana Irsova. Relative risk aversion: A meta-analysis. *EconS-tor Preprints*, (260586), 2022.
- [15] Christian Gollier and John W. Pratt. Risk vulnerability and the tempering effect of background risk. *Econometrica*, 64:1109–1124, 1996.
- [16] Francisco Gomes and Alexander Michaelides. Portfolio choice with internal habit formation: A life-cycle model with uninsurable labor income risk. *Review of Economic Dynamics*, 6(4):729–766, 2003.
- [17] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(4):263–291, 1979.
- [18] D. Kahneman and A. Tversky. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.
- [19] Miles S. Kimball. Precautionary savings in the small and in the large. *Econometrica*, pages 53–73, 1990.
- [20] Miles S. Kimball. Precautionary motives for holding assets. In *P. Newman, M. Milgate, and J. Falwell, eds.* MacMillan, London, 1992.
- [21] H. M. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, March 1952.
- [22] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

- [23] Lisa Morhaim and Ayşegül Yıldız Ulus. On history-dependent optimization models: A unified framework to analyze models with habits, satiation and optimal growth. *Journal of Mathematical Economics*, 105:102807, 2023.
- [24] Claus Munk. 1495 Preferences. In *Financial Asset Pricing Theory*. Oxford University Press, 04 2013.
- [25] Thorsten Pachur and David Kellen. Modeling gain-loss asymmetries in risky choice: The critical role of probability weighting. 2013.
- [26] John W. Pratt. Risk aversion in the small and in the large. *Econometrica*, pages 122–136, 1964.
- [27] Enrichetta Ravina. Habit persistence and keeping up with the joneses: Evidence from micro data. November 2007.
- [28] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, July 1948.
- [29] T. D. Stanley. Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15(3):131–150, 2001.

## 7. Apéndice

### 7.1. Principios básicos sobre funciones de utilidad

A modo de introducir los conceptos básicos discutidos en este trabajo al lector no especializado, se plantea aquí una síntesis del concepto de utilidad y preferencias siguiendo los lineamientos presentados en el capítulo 1 del libro "Microeconomic Theory" de Mas-Collel [22]. La utilidad es un concepto fundamental en la teoría económica que refleja las preferencias de un individuo en relación con las diferentes combinaciones de bienes y servicios disponibles para su consumo. Se define como una función de utilidad, denotada generalmente como  $U(x)$ , donde  $U$  es la función de utilidad del consumidor, y  $x$  es un vector que representa la cantidad y calidad de bienes y servicios consumidos por el individuo.

Asimismo toda función de utilidad debe cumplir con las siguientes propiedades:

1. **Ordenamiento de Preferencias:** La función de utilidad permite ordenar las preferencias del consumidor. Si dos cestas de bienes  $x_1$  y  $x_2$  son tales que  $U(x_1) > U(x_2)$ , entonces el consumidor prefiere la cesta  $x_1$  a la cesta  $x_2$ , es decir  $x_1 \succ x_2$ . Si  $U(x_1) = U(x_2)$ , el consumidor es indiferente entre las dos cestas ( $x_1 \sim x_2$ )
2. **Completitud de las Preferencias:** Para cualquier par de cestas de bienes  $x_1$  y  $x_2$ , el consumidor debe ser capaz de compararlas de una y solo una de las siguientes maneras:  $x_1 \succ x_2$  (i.e.,  $x_1$  es preferible a  $x_2$ ),  $x_2 \succ x_1$  (i.e.,  $x_2$  es preferible a  $x_1$ ), o  $x_1 \sim x_2$  (i.e., el individuo es indistinto ante cualquier alternativa).
3. **Monotonía:** Si todos los bienes en una cesta de bienes se incrementan en cantidad, la utilidad del consumidor no disminuye. Formalmente, si  $x_1 \geq x_2$  (esto significa que  $x_1$  contiene al menos la misma cantidad de todos los bienes que  $x_2$ ), entonces  $U(x_1) \geq U(x_2)$ .
4. **Transitividad:** Si un consumidor prefiere una cesta de bienes  $x_1$  a otra cesta  $x_2$  y también prefiere  $x_2$  a otra cesta  $x_3$ , entonces el consumidor debe preferir  $x_1$  a  $x_3$ .

Estas propiedades de la función de utilidad permiten a los economistas modelar y analizar las decisiones de consumo de los individuos de manera coherente y ayudan a derivar conclusiones importantes en la teoría económica, como la teoría de la demanda, la teoría del equilibrio del consumidor y el análisis del bienestar. En el particular caso que se aborda en este trabajo, los *bienes* (o canasta de bienes) se interpretan como instrumentos financieros (o portfolios de instrumentos); las propiedades discutidas en la siguiente sección se aplican sobre el marco general de funciones de utilidad como si de bienes convencionales se tratase.