



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

DOBLE GRAU DE MATEMÀTIQUES I ADMINISTRACIÓ I DIRECCIÓ D'EMPRESES

Treball final de grau

Estudi Predictiu de *Swaps* de Volatilitat Topada amb Algorismes d'Aprenentatge Automàtic

Autora: Laura Bartralot Rodríguez

Director: Dr. José Manuel Corcuera
Realitzat a: Departament de Matemàtiques
i Informàtica

Barcelona, 15 de gener de 2025

Abstract

This paper presents a methodology for pricing financial instruments such as capped volatility swaps using machine learning techniques. Its price, given as a function of volatility, is relegated to second place and brought the underlying asset's volatility into focus, essential part of the instrument. In order to predict the volatility and, consequently, establish a price, we examine distributional characteristics of the underlying. With this aim, we investigate market-implied moments and the implied volatility of the asset. Finally, using machine learning techniques and a Gaussian process regression we evaluate the predictive performance of the methods used and the dataset itself, thus establishing a model for both verification and prediction of prices.

Resum

Aquest treball desenvolupa una metodologia per a la valoració d'un instrument financer com són els *swaps* de volatilitat topada mitjançant tècniques d'aprenentatge automàtic. El preu d'aquests instruments, donat com a funció de la volatilitat, es deixa en segon pla i se centra l'atenció en la volatilitat de l'actiu subjacent, part essencial de l'instrument. Per tal de predir aquesta volatilitat i, en conseqüència, determinar un preu, s'estudien les característiques distribucionals dels actius subjacents. Amb aquest objectiu s'investiguen en concret els moments implícits del mercat i la volatilitat implícita dels actius. Finalment, amb mètodes d'aprenentatge automàtic i la regressió del procés Gaussià s'avalua la capacitat predictiva dels mètodes i de les dades, establint així un model tant de verificació com de predicció de preus.

Agraïments

Al meu tutor, José Manuel, per la valuosa orientació i seguiment al llarg de tot el procés.
A la meva família, per creure en mi incondicionalment.
A les amigues i amics, pels que veuré aquests anys amb un somriure.

Recordo quan li explicava a la meva àvia quina carrera estava estudiant i ella sempre responia amb un: “I això per a què serveix?”, al que jo contestava esclatant de riure. Uns quants anys més tard crec que sóc capaç de respondre i, encara millor, que ella entengui la resposta a una pregunta que, tanmateix, té una infinitat de respostes.

*“La incertidumbre es la única certeza que existe,
y saber cómo vivir con la inseguridad es la única seguridad”*

- John Allen Paulos

Índex

1	Introducció	1
2	Terminologia bàsica	3
2.1	Arbitratge	3
2.2	Posicions llarga i curta	3
2.3	Cartera o <i>portfolio</i>	3
3	Què és un <i>swap</i> de volatilitat?	4
3.1	Instruments financers derivats	4
3.2	Opcions financeres	4
3.3	<i>Swaps</i> o permutes financeres	5
3.3.1	La volatilitat	5
4	Processos estocàstics de temps continu	7
4.1	Moviment Brownià	7
4.2	Martingales	8
5	Model de Black-Scholes	10
5.1	Teorema del Límit Central	10
5.2	Moviment Brownià geomètric	10
5.3	La fórmula d'Itô	11
5.4	L'equació de Black-Scholes	13
5.5	La fórmula de Black-Scholes	14
5.6	Paritat <i>put-call</i>	15
5.7	Limitacions del model	15
6	Metodologia de recopilació de dades	16
6.1	Estimació de variables	17
6.1.1	Estimació de K	17
6.1.2	Estimació d' <i>IVOL</i>	19
7	Obtenció de variables d'entrada	20
7.1	Moments implícits del mercat	20
7.1.1	Simulació d' <i>inputs</i>	22
7.2	Volatilitat implícita	23
7.2.1	Simulació d' <i>inputs</i>	27

8	Models predictius	30
8.1	Regressió del Procés Gaussià	30
8.2	Aprenentatge automàtic basat en arbres	34
8.3	Bosc Aleatori	35
8.3.1	Bagging	35
8.3.2	CART	36
8.3.3	Algorisme	36
8.4	Potenciació del Gradient (<i>gradient boosting</i>)	37
8.4.1	Algorisme	38
8.5	Validació	38
8.5.1	Validació creuada (<i>cross-validation</i>)	38
8.5.2	<i>Walk-forward validation</i>	39
9	Avaluació	40
10	Conclusions	43

1 Introducció

El 6 de maig del 2010, els mercats financers dels Estats Units van viure un esdeveniment sense precedents, conegut com a **Flash Crash**. En qüestió de 30 minuts, l'índex *Dow Jones Industrial Average* (entre d'altres), més conegut simplement com a *Dow Jones* (de forma errònia, ja que hi ha altres índex *Dow Jones*) va partir una caiguda de quasi 1.000 punts, que suposaven un 9% del seu valor total. Aquesta caiguda va ser seguida per una recuperació immediata de quasi tots els punts que acabava de perdre, deixant als analistes i inversors perplexos amb el que acabava de succeir.

Tot i que inicialment es va sospitar que la causa havia estat un atac cibernètic, les investigacions posteriors van revelar que el succés va ser desencadenat per un conjunt de factors. La preocupació pel deute públic grec, les pràctiques de manipulació del mercat com el *spoofing* i els algorismes de *trading* d'alta freqüència (HFT), dissenyats per aprofitar diferències mínimes en preus dels mercats, que llancen ordres de compra i venda en mil·lisegons van crear un efecte dominó. Davant la caiguda inicial, aquests algorismes van reaccionar de manera agressiva i això va fer que la volatilitat augmentés, cosa que va provocar que altres sistemes s'activessin i així successivament.

El **Flash Crash** va posar en evidència els riscos inherents d'un mercat fortament automatitzat, on la interacció d'algorismes pot desencadenar successos de volatilitat extrema. També va posar de manifest la fragilitat del sistema financer i com, en qüestió de minuts, la volatilitat pot convertir un mercat aparentment estable en un caos total. Per últim, aquest incident va recalcar la necessitat de l'existència d'un equilibri entre les operacions i supervisió humana i les operacions automatitzades o *trading* automatitzat.

La gestió dels riscos i l'especulació han portat als mercats i als seus agents a desenvolupar instruments financers que permetin cobrir unes necessitats latents per a tota mena d'empreses. Entre aquests hi trobem els *swaps* de volatilitat topada, que destaquen com a eines per gestionar l'exposició a la incertesa dels preus d'actius i per a especular. Paral·lelament al desenvolupament d'instruments, els acadèmics s'han interessat a crear eines, teories i fórmules per a quantificar i estimar de la millor forma possible tots els elements que componen el mercat financer.

Com es veurà més endavant, la formalització matemàtica, duta a terme per Robert Weiner l'any 1923, del moviment Brownià, descobert per Robert Brown l'any 1828, va suposar un gran avenç per descriure els moviments del mercat, que deuen el seu comportament a un component estrictament aleatori. Aquesta aleatorietat intrínseca dels mercats no és res més que un reflex de la naturalesa de l'ésser humà i dels seus comportaments, a vegades irracionals, que tenen un paper notable en els moviments dels preus. Això és exactament el que va formular Louis Bachelier l'any 1900 en la seva tesi doctoral, on va introduir la idea que les fluctuacions dels preus dels actius financers podien modelar-se amb un procés aleatori continu.

El moviment Brownià, juntament amb altres resultats, van establir les bases per què Samuelson en primera instància i Black, Scholes i Merton, més endavant, extraguessin un resultat, en forma de fórmula senzilla, per calcular els preus d'un instrument molt conegut al mercat, les opcions financeres.

Més recentment, i degut a l'augment de la capacitat computacional dels ordinadors, hi ha hagut grans avenços en un camp com és l'aprenentatge automàtic o *machine learning*. L'objectiu d'aquest és el desenvolupament de tècniques i models que permetin que els ordinadors aprenguin a partir de les dades i, d'aquesta manera, puguin fer projeccions per al futur.

L'objectiu d'aquest treball és la implementació de tècniques d'aprenentatge automàtic per a la valoració dels *swaps* de volatilitat topada. Per això caldrà establir les bases teòriques per determinar els preus dels *swaps* de volatilitat topada, així com els preus de les opcions financeres. Seguidament, es presentaran els models o tècniques d'aprenentatge automàtic que es faran servir per avaluar la capacitat predictiva. Finalment, es presentaran els resultats obtinguts en implementar els models i n'avaluarem el seu rendiment.

2 Terminologia bàsica

En aquesta secció es defineixen breument conceptes necessaris per al desenvolupament i comprensió del treball.

2.1 Arbitratge

Estratègia o pràctica que es duu a terme als mercats financers basada en la capacitat de generar beneficis sense incórrer en riscos. Això és possible, ja que en diferents mercats, un mateix actiu financer pot tenir diferents preus. L'estratègia, doncs, és comprar en el mercat de preus baixos i vendre en el de preus més elevats, obtenint un benefici de la diferència. Aquesta diferència es deu a les ineficiències presents als mercats. Com més eficient és un mercat, més difícil és realitzar l'arbitratge.

Durant tot el treball ens referirem a condicions on no s'admet l'arbitratge, això implica que no és possible obtenir beneficis sense risc o el que és el mateix, sense invertir un capital inicial.

2.2 Posicions llarga i curta

Al mercat de valors existeixen dues formes d'exposar-se a un actiu subjacent. Aquestes es basen en l'expectativa que es té de l'evolució del preu de l'actiu. Si es té una expectativa alcista, és a dir, que el preu de l'actiu augmentarà, l'estratègia evident és comprar per vendre en el futur i obtenir benefici de la diferència. A aquesta se l'anomena posició llarga.

En contrast, quan l'expectativa és baixista, això és, que el preu de l'actiu baixarà, es pren una posició curta. L'estratègia en aquest cas consisteix a agafar en préstec l'actiu per vendre'l, i passat un temps en què el preu hagi baixat, recomprar l'actiu per tornar-lo.

En ambdós casos, si el preu de l'actiu fluctua en direcció contrària a l'expectativa, s'incorraran en pèrdues.

2.3 Cartera o *portfolio*

Una cartera és qualsevol combinació d'actius financers com accions, bons, etc. El seu valor és la suma dels valors de cada actiu.

3 Què és un *swap* de volatilitat?

Abans de definir què és un *swap* de volatilitat, cal entendre quin tipus d'actiu és i quin lloc ocupa en la classificació d'actius en un mercat financer.

3.1 Instruments financers derivats

Un mercat financer és un espai físic o virtual en què els agents econòmics compren i venen actius i en el que es determinen els seus preus. Dins d'aquest, podem trobar un gran ventall d'actius que es poden intercanviar. Entre els actius més famosos hi trobem les accions, els bons i les obligacions.

Tanmateix, hi ha molts altres instruments financers. Per exemple, existeixen instruments que basen el seu valor en el preu d'un altre actiu. Aquests s'anomenen derivats financers o senzillament derivats. L'actiu en què es basa el valor dels derivats s'anomena **actiu subjacent** i aquests solen ser accions, bons, matèries primeres, tipus d'interès, índexs borsaris o monedes (*currencies*).

Aquests instruments representen un contracte en què dues parts acorden una transacció basada en el valor de l'actiu subjacent en un moment del temps determinat. Dins d'aquest tipus d'instruments hi trobem els futurs, els *forwards*, les opcions i les permutes (*swaps*).

3.2 Opcions financeres

Una opció financera és un contracte entre parts que dona el dret, però no l'obligació, a comprar (*call*) o vendre (*put*) béns o actius a un preu predeterminat (*Strike* o preu d'exercici) fins a una data determinada (venciment). El venedor de l'opció té l'obligació de vendre o comprar en cas que el comprador vulgui exercir el seu dret.

Les opcions més senzilles s'anomenen opcions 'vainilla' que consisteixen en contractes bàsics d'opcions *call* i *put*. Dins d'aquestes, hi ha les opcions europees i les americanes. La diferència entre aquestes és que en el cas de les primeres, les opcions només poden ser exercides al venciment. En canvi, les opcions americanes poden ser exercides en qualsevol moment al llarg de la vida de l'opció, fins al venciment. D'ara en endavant considerarem només les opcions europees 'vainilla'.

El benefici d'aquestes opcions ve determinat per la diferència entre el preu de l'actiu subjacent al venciment i el preu d'exercici, en el cas d'un *call*

$$(S_T - K)_+ = \max(0, S_T - K)$$

i en el cas d'un *put*

$$(K - S_T)_+ = \max(0, K - S_T)$$

És clar que el comprador d'una opció *call* espera que el preu S_T de l'actiu al venciment sigui superior a K , aquesta és una estratègia *alcista*. En contrast, el comprador d'una opció *put* espera que S_T sigui inferior a K . En aquest cas, l'estratègia és *baixista*.

Així mateix, el venedor d'una opció *call* adopta una estratègia no *baixista* i el venedor d'una opció *put* una no *alcista*.

Un dels problemes clàssics és com determinar aquest preu K , aquest s'anomena *The pricing problem* o problema de la fixació de preus.

3.3 *Swaps* o permutes financeres

Els *swaps* són contractes que consisteixen en intercanvis de fluxos monetaris entre dues parts en dates futures. Els *swaps* més comuns al mercat són de taxes d'interès, de moneda (*currency*), d'inflació, de volatilitat, de variància i de mercaderies. La majoria d'aquests últims involucren el petroli, de forma que moltes companyies aèries i ferroviàries acorden *swaps* de productes petrolers per tal de garantir costos més baixos a llarg termini.

Un *swap* de volatilitat és un contracte entre dues parts, en un mercat extraborsari (Over The Counter), basat en la volatilitat realitzada d'un actiu subjacent. Al final del contracte es liquida en funció de la diferència entre la volatilitat realitzada σ_R i un nivell de volatilitat fixat a l'inici del contracte (Strike o K).

El preu Strike, K o preu d'exercici és un preu que es determina a l'inici del contracte. En el cas dels *swaps* de volatilitat, aquest preu acabarà determinant si el comprador del contracte obtindrà beneficis o pèrdues i reflecteix les expectatives de volatilitat del mercat a l'inici del contracte. Més endavant s'aclarirà com es determina aquest preu.

Per tant, aquest tipus de contracte és interessant tant per la cobertura de possibles riscos com per especular únicament sobre el moviment de volatilitat d'un actiu, sense negociar directament aquest actiu, tal com descriuen Demeterfi *et al.* (1999) [11].

Hi ha una gran varietat d'actius subjacents sobre els quals es poden efectuar operacions financeres i poden ser tant accions, bons, divises o índexs borsaris. En el present treball, els actius subjacents amb què treballarem seran l'índex **Standard & Poor's 500 (S&P 500)** i accions individuals de **JPMorgan Chase & Co (JPM)**, ambdós cotitzant a la borsa de Nova York.

3.3.1 La volatilitat

La volatilitat és una mesura de freqüència i intensitat dels canvis del preu d'un actiu. Es fa servir per a quantificar el risc d'un instrument. Una volatilitat alta indica una ràpida i forta fluctuació de preus, el que pot comportar majors oportunitats de benefici i pèrdua. En canvi, una volatilitat baixa indica un instrument estable, de baix risc, atractiu per a inversors conservadors.

Generalment, es calcula la volatilitat usant la desviació típica.

La volatilitat realitzada és una mesura *ex-post*, és a dir, mesura la volatilitat observada al mercat durant un temps determinat. Es basa en els retorns diaris de l'actiu subjacent i el seu càlcul es fa prenent l'arrel quadrada de la variància realitzada. En primer lloc, es calculen els log-retorns per a cada dia: $\log(S_i) - \log(S_{i-1})$. Així doncs, calculant la suma dels retorns quadrats al llarg del període T que consta de N dies, obtenim la variància realitzada

$$\sigma_R^2 = \sum_{i=1}^N \left(\log \left[\frac{S_i}{S_{i-1}} \right] \right)^2$$

Finalment, s'anualitza o escala aquesta mesura amb una constant $\frac{252}{N}$ i es pren l'arrel quadrada, on 252 són els dies de *trading* al llarg d'un any. Finalment, obtenim que la volatilitat realitzada ve donada per la fórmula:

$$\sigma_R = 100 \times \sqrt{\frac{252}{N} \sum_{i=1}^N \left(\log \left[\frac{S_i}{S_{i-1}} \right] \right)^2}$$

On S_i és el preu de tancament de l'actiu subjacent al dia i , i N els dies de *trading* entre l'inici del contracte i el final en T .

Per tal de mitigar el risc d'aquests de volatilitat, sovint s'introdueix un topall (*cap*) a la volatilitat realitzada. Normalment, aquest topall se situa a 2,5 vegades la volatilitat fixada a l'inici del contracte (*Strike*). Aquests contractes es coneixen com a *swaps* de volatilitat topada o *capped volatility swaps*.

El benefici d'un instrument com aquest ve donat per la següent expressió, tal com descriu Flavell (2010) [12]

$$b = \text{nocional} \times [\min(\text{cap}, \text{volatilitat realitzada}) - \text{Strike}]. \quad (3.1)$$

4 Processos estocàstics de temps continu

Definició 4.1. Un procés estocàstic és una família de variables aleatòries reals $(X_t)_{t \geq 0}$ definida en un espai de probabilitats $(\Omega, \mathcal{F}, \mathbb{P})$.

Definició 4.2. Una filtració $(\mathcal{F}_t)_{t \geq 0}$ és una seqüència creixent de sub- σ -àlgebres de $\{\mathcal{F}_t, t \in [0, T]\}$ complint que $\mathcal{F}_s \subseteq \mathcal{F}_t$ per a tot $s \leq t$. $\mathcal{F}_t$ representa la col·lecció de tots els esdeveniments observables fins al moment t . Prenem $\mathcal{F}_0 = \{\emptyset, \Omega\}$ i $\mathcal{F}_T = \mathcal{F}$.

Definició 4.3. Un procés estocàstic $(X_t)_{t \geq 0}$ definit a l'espai de probabilitats $(\Omega, \mathcal{F}, \mathbb{P})$ es diu que és adaptat a una filtració $(\mathcal{F}_t)_{t \geq 0}$ si per a tot $t \geq 0$, X_t és $\mathcal{F}_t$ -mesurable.

Observació 4.4. Podem construir la filtració generada pel procés $(X_t)_{t \geq 0}$ i escriure $\mathcal{F}_t = \mathcal{N} \vee \sigma(X_s, 0 \leq s \leq t) = \bar{\sigma}(X_s, 0 \leq s \leq t)$ on $\mathcal{N} \subseteq \mathcal{F}_0$ i $\mathcal{N}$ és la col·lecció dels conjunts negligibles (conjunts amb probabilitat zero).

4.1 Moviment Brownià

El moviment Brownià descriu el moviment irregular i aleatori que segueixen algunes partícules microscòpiques en un fluid. La història del moviment Brownià es remunta fins al 1828 i rep aquest nom en honor a Robert Brown, biòleg i botànic, que mentre mirava a través d'un microscopi les partícules atrapades en les cavitats d'un gra de pol·len a l'aigua, va observar que aquestes es movien a través del líquid de forma constant i irregular. Replicant-ho amb partícules de pols, va poder descartar que el moviment fos degut al fet que les partícules de pol·len estiguessin 'vives'.

Al 1900, Louis Bachelier va considerar el moviment Brownià com a un possible model per a preus del mercat de valors [1]. Però no va ser fins al 1923, que Norbert Wiener va definir i construir rigorosament el moviment Brownià per primera vegada. El procés estocàstic que se'n resulta és sovint conegut com a procés de Wiener en el seu honor.

Definició 4.5. Diem que $(X_t)_{t \geq 0}$ és un procés amb increments independents si per tot $0 \leq t_1 \leq \dots \leq t_n$, es té que $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ són variables aleatòries independents.

Observació 4.6. De forma equivalent podem dir que $(X_t)_{t \geq 0}$ és un procés amb increments independents si per tot $0 \leq s \leq t$, $X_t - X_s$ és independent de $\sigma(X_u, 0 \leq u \leq s) = \mathcal{F}_s$.

Definició 4.7. Un moviment Brownià és un procés estocàstic a temps continu amb increments independents, estacionaris i amb trajectòries contínues. És a dir,

- $\mathbb{P}$ -c.s. $s \rightarrow X_s(\omega)$ és continu
- $s \leq t$, $X_t - X_s$ és independent de $\sigma(X_u, 0 \leq u \leq s) = \mathcal{F}_s$
- $s \leq t$, $X_t - X_s \sim X_{t-s} - X_0$

Teorema 4.8. Si $(X_t)_{t \geq 0}$ és un moviment Brownià aleshores

$$X_t - X_0 \sim \mathcal{N}(rt, \sigma^2 t) \quad (4.1)$$

La demostració es pot trobar al capítol 2 de Karatzas i Shreve (1988) [17].

Definició 4.9. Diem que un moviment Brownià és estàndard si $X_0 = 0$ $\mathbb{P}$ c.s., $r = 0$ i $\sigma^2 = 1$.

A partir d'ara, ens referirem al **moviment Brownià estàndard** simplement com a **moviment Brownià**.

Observació 4.10. Veiem que $X_t - X_s \sim \mathcal{N}(0, t - s)$ per a tot $0 \leq s \leq t$.

Aquest procés serà de vital importància per a la construcció del model de Black-Scholes.

4.2 Martingales

Sigui $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ un espai de probabilitat filtrat, on $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ és una filtració.

Definició 4.11. Una seqüència de variables aleatòries $(X_t)_{t \geq 0}$ és una martingala si satisfà:

1. Per tot $t \geq 0$, X_t és $\mathcal{F}_t$ -mesurable, és a dir, X_t és conegut en t .
2. Per tot $t \geq 0$, X_t és integrable, és a dir, $\mathbb{E}[|X_t|] < \infty$.
3. Per tot $s \leq t$ es compleix que el valor esperat de X_t donat $\mathcal{F}_s$ és X_s , és a dir,

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s \quad \text{c.s. } \mathbb{P}$$

Proposició 4.12. Si (X_t) és un moviment Brownià (estàndard) aleshores (X_t) és una martingala.

Demostració. Sigui $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$, la demostració és immediata considerant $\mathbb{E}(X_t | \mathcal{F}_s) = \mathbb{E}(X_t - X_s + X_s | \mathcal{F}_s)$ i aplicant l'observació 4.10.

Definició 4.13. Diem que dues mesures de probabilitat $\mathbb{P}$ i $\mathbb{P}^*$ en el mateix espai mesurable $(\Omega, \mathcal{F})$ són equivalents si tenen els mateixos conjunts de probabilitat zero. És a dir, si $\mathcal{N}$ denota el conjunt d'esdeveniments amb probabilitat zero, aleshores $\mathcal{N}_{\mathbb{P}} = \mathcal{N}_{\mathbb{P}^*}$. Escrivim $\mathbb{P} \sim \mathbb{P}^*$.

Definició 4.14. Diem que una mesura $\mathcal{Q}$ o $\mathbb{P}^*$ és una mesura neutral al risc en un espai de probabilitats $(\Omega, \mathcal{F}, \mathbb{P})$ filtrat si és equivalent a $\mathbb{P}$ ($\mathcal{Q} \sim \mathbb{P}$) i es compleix que els preus descomptats dels actius sota aquesta mesura són martingales ($\mathcal{Q}$ -martingala).

Una mesura neutral al risc ($\mathcal{Q}$ o $\mathbb{P}^*$) és una forma d'ajustar les probabilitats reals (amb mesura $\mathbb{P}$) de manera que es descriu un escenari on els actius tenen el mateix creixement esperat que la taxa lliure de risc r . Aquesta mesura neutral al risc és una mesura de probabilitat sota la qual el valor esperat de qualsevol actiu descomptat per la taxa lliure de risc segueix un procés de martingala, (per això també es coneix com a mesura d'equivalència martingala). És a dir, diem que un actiu S_t descomptat de la taxa lliure de risc és una martingala sota la mesura $\mathcal{Q}$ si es compleix que:

$$\mathbb{E}^{\mathcal{Q}}[F D_t \cdot S_T | \mathcal{F}_t] = \mathbb{E}^{\mathcal{Q}}\left[\frac{S_T}{e^{r(T-t)}} \middle| \mathcal{F}_t\right] = \frac{S_t}{e^{r(t-t)}} = S_t$$

On $\mathbb{E}^{\mathcal{Q}}[\cdot]$ denota l'esperança sota la mesura neutral al risc $\mathcal{Q}$. Aleshores, per determinar el valor d'un actiu en el moment actual t tenim que

$$S_t = FD_t \cdot \mathbb{E}^{\mathcal{Q}}[S_T | \mathcal{F}_t] \quad (4.2)$$

De forma anàloga, tenim que sota la mesura $\mathcal{Q}$, podem expressar el preu d'una opció com els pagaments futurs o benefici esperat que n'obtindrem descomptat al present seguint el següent resultat. La Teoria de valoració de preus per arbitratge (o Teoria de martingales) diu que en un mercat sense oportunitats d'arbitratge, el preu d'un actiu derivat ha d'ajustar-se de forma que el seu valor descomptat sigui una martingala sota la mesura $\mathcal{Q}$ segons Ross (1976) [23] i Harrison i Pliska (1981) [15]. Sigui $(S_T - K)_+$ el benefici que s'obtindria en T d'una opció *call*, aleshores

$$Preu_t = FD_t \cdot \mathbb{E}^{\mathcal{Q}}[(S_T - K)_+ | \mathcal{F}_t] \quad (4.3)$$

El benefici d'un *swap* de volatilitat topada ve donat per $(\min(cap, \sigma_R) - K)$, i aleshores el preu de l'instrument en un moment determinat t ve donat per

$$Preu_t = FD_t \times \mathbb{E}_t^{\mathcal{Q}}[\min(cap, \sigma_R) - K] \quad (4.4)$$

Aquest $FD_t = \exp(-r(T-t))$ és un factor de descompte determinat per la taxa d'interès r lliure de risc al mercat i la maduresa T del contracte.

El preu *Strike* es determina de tal manera que hi ha paritat a l'inici del contracte, és a dir, que l'*Strike* és el valor esperat a l'inici del contracte ($t = 0$) de la volatilitat de l'actiu subjacent durant la vida del contracte. Per tant, $K = \mathbb{E}_0^{\mathcal{Q}}[\min(cap, \sigma_R)]$, ja que

$$\begin{aligned} Preu_0 &= FD_0 \times \mathbb{E}_0^{\mathcal{Q}}[\min(cap, \sigma_R) - K] \\ &= FD_0 \times (\mathbb{E}_0^{\mathcal{Q}}[\min(cap, \sigma_R)] - \mathbb{E}_0^{\mathcal{Q}}[K]) \\ &= FD_0 \times (\mathbb{E}_0^{\mathcal{Q}}[\min(cap, \sigma_R)] - K) = 0 \end{aligned}$$

D'on obtenim la tercera igualtat degut a que K és constant.

La naturalesa no lineal de l'expressió, tant del benefici b en (3.1) com del *Preu*, fa que trobar una expressió tancada i senzilla resulti complex. D'aquesta forma ens centrarem en desenvolupar un mètode basat en les dades del mercat per construir un model de fixació de preus.

5 Model de Black-Scholes

A principis del anys 70, Fischer Black, Myron Scholes i Robert C. Merton van fer un gran avenç en la valoració d'opcions desenvolupant el que es coneix com el model de Black-Scholes. La importància d'aquest model va fer-se palesa quan al 1997 Scholes i Merton van ser premiats amb el Premi Nobel d'economia. Black, que va morir dos anys abans, no va veure com la seva contribució era tant reconeguda.

Aquest model estudia la dinàmica d'un mercat financer que conté instruments d'inversió derivats que es plasma en el que es coneix com l'equació de Black-Scholes. A partir d'aquesta, se'n dedueix la fórmula de Black-Scholes, que dona una estimació teòrica del preu de les opcions d'estil europees (*vanilla european options*), d'una forma relativament senzilla i amb un nombre tancat de variables.

5.1 Teorema del Límit Central

Abans d'introduir el model, s'enuncia aquest teorema que es mencionarà per entendre la construcció del model. Aquest teorema estableix la convergència en llei cap a la distribució normal de la successió de sumes de variables aleatòries independents. El fet sorprenent és que el límit és independent de la llei de les variables de la successió.

Teorema 5.1. (Teorema del Límit Central de Lévy-Lindeberg) *Sigui $(X_n)_{n \geq 1}$ una successió de variables aleatòries i.i.d i de quadrat integrable, amb variància $\sigma^2 > 0$ i mitjana m . Si $S_n = X_1 + \dots + X_n$, es compleix que*

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{\text{Llei}} \mathcal{N}(0, 1)$$

Seguint el capítol 6 de Sanz-Solé (1999) [24].

5.2 Moviment Brownià geomètric

El model de Black-Scholes suposa que el mercat financer consta de dos tipus d'actius, un actiu sense risc S^0 que evoluciona com

$$dS_t^0 = rS_t^0 dt \quad 0 \leq t \leq T, \quad S_0^0 = 1$$

per tant,

$$S_t^0 = e^{rt}$$

on $t \geq 0$ i r és una constant no negativa, i un actiu amb risc $(S_t)_{0 \leq t \leq T}$ evoluciona de la següent forma:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad t \neq 0 \quad (5.1)$$

on (B_t) és un moviment Brownià.

La idea darrere d'aquesta equació diferencial es descriu a continuació tal com exposen Madan i Schoutens (2016) [20].

El model estudia l'evolució del preu d'un actiu del mercat de valors, sigui $S = \{S_t, t \geq 0\}$. Considerem l'evolució d' S durant un període de temps des del present t fins a $t + \Delta t$. Escrivim ΔS_t al canvi $S_{t+\Delta t} - S_t$, el retorn relatiu durant aquest període és $\frac{\Delta S_t}{S_t}$. És raonable esperar que aquest retorn es pugui descompondre en dues components, una part *sistemàtica* i una part *aleatòria*.

Analitzem primer la part sistemàtica. Assumim que el retorn esperat durant un període serà proporcional a la longitud d'aquest període. Sigui $[t, t + \Delta t]$ el període, de longitud Δt . Per tant, el retorn esperat d'aquest període, proporcional a la longitud, ve donat per $\mu \Delta t$, on μ és un paràmetre que representa la taxa mitjana dels retorns de l'actiu.

En segon lloc, donat que el preu d'un actiu del mercat de valors fluctua estocàsticament, una assumpció raonable és que la variància del retorn també sigui proporcional a la longitud de l'interval $[t, t + \Delta t]$. Tenint en compte el Teorema del Límit Central i veient el retorn de l'actiu com la suma de molts efectes aleatoris, té sentit modelar la part aleatòria dels retorns amb una variable aleatòria distribuïda Normalment, amb variància $\sigma^2 \Delta t$, on σ descriu l'efecte o pes que té aquesta variable (o com és de volàtil el retorn); a σ l'anomenem volatilitat de l'actiu.

Unint les dues parts i considerant que si B_t és un moviment Brownià aleshores tenim que $\sigma \Delta B_t = \sigma(B_{t+\Delta t} - B_t)$ té una distribució normal amb la variància requerida $\sigma^2 \Delta t$ pel teorema 4.8, podem escriure

$$\Delta S_t = S_t(\mu \Delta t + \sigma \Delta B_t), \quad S_0 > 0$$

Prenent $\Delta t \rightarrow 0$, tenim l'equació diferencial estocàstica (5.1).

5.3 La fórmula d'Itô

Definició 5.2. Un procés d'Itô en $[0, T]$ és un procés estocàstic X que verifica

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dB_s$$

on

- X_0 és $\mathcal{F}_0$ -mesurable
- (K_t) i (H_t) són $(\mathcal{F}_t)$ -adaptats
- $\int_0^T (|K_s| + |H_s|^2) ds < \infty$ $\mathbb{P}$ -c.s..

Una pràctica generalment acceptada és fer servir la notació diferencial

$$dX_t = K_t dt + H_t dB_t$$

Definició 5.3. Donada una aplicació $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, diem que és (globalment) Lipschitz respecte x en U si existeix una constant $K > 0$ tal que

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|, \quad \forall x_1, x_2 \in U$$

Teorema 5.4. Donada una equació diferencial

$$dX_t = f(t, X_t)dt + g(t, X_t)dB_t \quad (5.2)$$

tal que $f(t, x)$ i $g(t, x)$ són Lipschitz respecte x , amb creixement lineal i continues respecte t , de manera que existeix $K > 0$ tal que

$$|f(t, x)| + |g(t, x)| \leq K(1 + |x|) \quad \forall x \in \mathbb{R}, t \in [0, T]$$

Aleshores l'equació diferencial té una única solució. (Seguint Capinski, Kopp i Traple (2012) [8])

Teorema 5.5. Sigui S_t un procés d'Itô del tipus

$$dS_t = f(t, S_t)dt + g(t, S_t)dB_t,$$

amb f i g funcions tals que existeix solució i és única. Aleshores per una funció $V_t = V(t, S_t) \in C^{1,2}$ se satisfà la següent fórmula

$$dV_t = \left(g(t, S_t) \frac{\partial V_t}{\partial S_t} dB_t \right) + \left(\frac{\partial V_t}{\partial t} + f(t, S_t) \frac{\partial V_t}{\partial S_t} + \frac{g(t, S_t)^2}{2} \frac{\partial^2 V_t}{\partial S_t^2} \right) dt \quad (5.3)$$

o alternativament

$$\begin{aligned} V(t, S_t) &= V(0, S_0) + \int_0^t V_t(s, S_s) ds + \int_0^t V_x(s, S_s) f(s, S_s) ds \\ &\quad + \int_0^t V_x(s, S_s) g(s, S_s) dB_s \\ &\quad + \frac{1}{2} \int_0^t V_{xx}(s, S_s) g(s, S_s)^2 ds \end{aligned}$$

on $V_x(t, S_t)$ denota la derivada de V respecte S_t i $V_t(t, S_t)$ respecte t .

Aplicant la fórmula de Itô a l'equació (5.1) prenent $V(t, S_t) = \log(S_t)$ tenim

$$\frac{\partial V_t}{\partial S_t} = \frac{1}{S_t}, \quad \frac{\partial^2 V_t}{\partial S_t^2} = -\frac{1}{S_t^2}, \quad \frac{\partial V_t}{\partial t} = 0$$

aleshores

$$dV_t = \left(\sigma S_t \frac{1}{S_t} dB_t \right) + \left(\mu S_t \frac{1}{S_t} - \frac{\sigma^2}{2} S_t^2 \frac{1}{S_t^2} \right) dt = \sigma dB_t + \left(\mu - \frac{\sigma^2}{2} \right) dt$$

i, per tant,

$$\log S_t = \log S_0 + \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t$$

Finalment obtenim que

$$S_t = S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right\} \quad (5.4)$$

Aquesta funció es coneix com a moviment Brownià geomètric. Més generalment

$$S_t = S_u \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) (t - u) + \sigma (B_t - B_u) \right\} \quad (5.5)$$

i, per tant, $\log(S_t)$ és un moviment Brownià, no necessàriament estàndard.

5.4 L'equació de Black-Scholes

A partir dels resultats anteriors es pot obtenir l'equació de Black-Scholes. Aquest model es basa en l'existència d'un actiu sense risc i un amb risc, que segueix un moviment Brownià geomètric. Les hipòtesis fonamentals del model sobre els actius són:

1. No es paguen dividends
2. La taxa d'interès lliure de risc és constant
3. Es presumeix que la volatilitat és constant
4. S_t segueix un procés de moviment Brownià geomètric.

I les hipòtesis sobre el mercat són:

1. No hi ha oportunitat d'arbitratge sense risc
2. No hi ha costos de transacció
3. Possibilitat de realitzar vendes en curt (*short-selling*) sense restriccions

Amb aquestes condicions, els autors van obtenir l'equació per al preu d'una opció de la següent manera.

Es construeix una cartera que consta de x unitats d'un actiu amb preu S_t i una posició curta d'una opció d'estil europeu de valor $V(t, S_t)$. El valor de la cartera és:

$$P(t, S_t) = xS_t - V(t, S_t)$$

Prenent diferencials

$$dP(t, S_t) = x dS_t - dV_t = x(\mu S_t dt + \sigma S_t dB_t) - dV_t$$

on $V(t, S_t)$ satisfà l'equació (5.3). Prenem $x = \frac{\partial V_t}{\partial S_t}$ i aleshores

$$dP_t = -\left(\frac{\partial V_t}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V_t}{\partial S_t^2}\right) dt$$

L'assumpció de no arbitratge implica que qualsevol cartera lliure de risc té el mateix retorn i es comporta de la següent forma:

$$dP_t = rP_t dt$$

ara només cal igualar les dues equacions anteriors

$$rP_t dt = -\left(\frac{\partial V_t}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V_t}{\partial S_t^2}\right) dt$$

i donat que $P_t = xS_t - V_t$ i $x = \frac{\partial V_t}{\partial S_t}$ tenim que

$$r\left(\frac{\partial V_t}{\partial S_t} S_t - V_t\right) = -\left(\frac{\partial V_t}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V_t}{\partial S_t^2}\right)$$

Amb tot, obtenim que el preu d'una opció satisfà la següent equació

$$\frac{\partial V_t}{\partial t} + rS_t \frac{\partial V_t}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} = rV_t \quad (5.6)$$

coneguda com l'**equació de Black-Scholes**.

5.5 La fórmula de Black-Scholes

A partir de l'equació anterior (5.6) i seguint Black i Scholes (1973) [3] s'obtenen les fórmules de Black-Scholes que permeten calcular el preu de les opcions europees. En el cas d'un *call* es té

$$C_t = S_t \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-) \quad (5.7)$$

i el d'un *put*

$$P_t = K e^{-r(T-t)} \Phi(-d_-) - S_t \Phi(-d_+) \quad (5.8)$$

on $\Phi(x)$ és la funció de distribució normal estàndard definida com $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz$, S_t és el preu de l'actiu subjacent en el moment t , K el preu *Strike* i on

$$d_{\pm} = \frac{\ln\left(\frac{S_t}{K}\right) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

Com s'ha comentat anteriorment, $\log(S_t)$ és un moviment Brownià i aleshores, per definició i per (5.5) tenim que

$$\log\left(\frac{S_t}{S_{t-1}}\right) = \left(\mu - \frac{\sigma^2}{2}\right) + \sigma(B_t - B_{t-1})$$

on $B_t - B_{t-1} \sim \mathcal{N}(0, 1)$, aleshores per la propietat de les distribucions normals que diu que si tenim $X \sim \mathcal{N}(\mu, \sigma^2)$ podem convertir-la en una distribució normal estàndard $Z \sim \mathcal{N}(0, 1)$ fent servir $X = \mu + \sigma Z$ tenim que

$$\log\left(\frac{S_t}{S_{t-1}}\right) \sim \mathcal{N}\left(\mu - \frac{\sigma^2}{2}, \sigma^2\right)$$

Podem observar que els log-retorns de l'índex **S&P 500** es distribueixen de la següent forma.

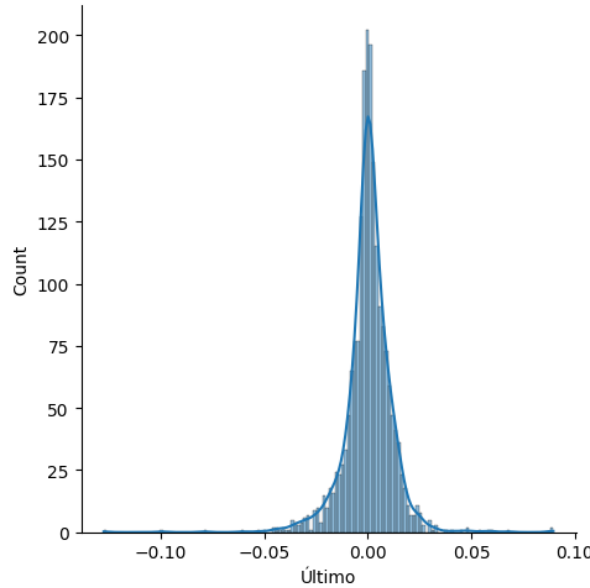


Figura 1: Log-retorns del **S&P 500** de l'any 2019 al 2022

5.6 Paritat *put-call*

La paritat *put-call* indica la relació existent entre el preu d'una opció de compra *call* i una opció de venda *put* en absència d'oportunitats d'arbitratge i amb el mateix preu d'exercici K , sempre que el venciment sigui el mateix per ambdues opcions. La determinació de la paritat només funciona amb opcions de compra i venda europees, ja que l'opció només pot exercir-se a la data de venciment.

Les opcions americanes poden exercir-se en qualsevol moment abans del venciment, cosa que dificulta la determinació de la paritat.

Suposem doncs un actiu, de preu S_t en el moment t , i que tenim una cartera formada per la compra d'un *call* i la venda d'un *put*. El benefici és:

$$(S_T - K)_+ - (K - S_T)_+ = S_T - K$$

Però aquest benefici pot ser replicat comprant una unitat de l'actiu i un títol que pagui exactament K en el moment T . Aquest títol podria ser l'anomenat bo cupó zero. Un bo cupó zero és un bo o títol emès per una entitat que no paga interessos durant la seva vida, sigui T , sinó que en el moment en què s'emet es ven per un preu inferior al seu valor nominal (sigui K) amb descompte i al venciment s'amortitza íntegrament pel seu valor nominal. Aquest descompte es pot entendre com el 'preu del diner' i com hem fet anteriorment, el definirem com $FD_t = e^{-r(T-t)}$.

Aleshores, si dues carteres tenen el mateix benefici, en un context en què no es permet l'arbitratge, el seu valor ha de ser idèntic per a tot t . D'aquesta manera la relació entre una opció *call* i una opció *put* és:

$$C_t - P_t = S_t - Ke^{-r(T-t)} \quad (5.9)$$

on K és el preu d'exercici d'ambdues opcions.

5.7 Limitacions del model

Tot i el gran avenç en matèria de valoració d'opcions que va suposar el model, cap és perfecte, i el de Black-Scholes no n'és una excepció. Les limitacions principals del model són:

- Tant la taxa d'interès com la volatilitat s'assumeixen constants al llarg de la duració de l'opció. En realitat cap de les dues sol ser-ho.
- S'assumeix que no hi ha costos de transacció al mercat, com els impostos, i s'ignora també el risc de liquiditat (possibilitat de no disposar dels diners en el termini acordat o necessari).
- Només és vàlid per opcions europees.
- S'assumeix que els preus dels actius segueixen una distribució log-normal. En realitat aquests presenten una lleugera asimetria.
- S'assumeix que no es paguen dividends.
- Altres assumpcions operacionals del model, com que no està permès l'arbitratge.

6 Metodologia de recopilació de dades

En aquesta secció es descriu la metodologia emprada en la recopilació de dades així com la justificació de l'elecció d'aquesta.

En l'expressió que proporciona el preu d'un *swap* en (4.4), s'observa que, en un moment donat t , les variables K i FD_t són fixes i, per tant, per estimar el preu d'un contracte, cal fer una estimació de la volatilitat realitzada σ_R .

Seguint [16] es proposa una fórmula alternativa que, en cada moment t i suposant conegut $Preu_t$, separi la part de volatilitat realitzada fins a t (coneguda) de la part encara no realitzada, és a dir, la volatilitat corresponent a $T - t$ (desconeguda).

Anomenarem $AcVOL$ a la volatilitat efectivament realitzada fins a t i $IVOL$ a la part encara no realitzada. D'aquesta forma la fórmula (4.4) queda de la següent manera

$$Preu_t = FD_t \times \left(\sqrt{IVOL_t^2 \times (1 - Pes_t) + AcVOL_t^2 \times Pes_t} - K \right) \quad (6.1)$$

on la variable Pes reflecteix la proporció de temps que ha passat des de l'inici del contracte fins al moment t respecte la duració total T . En el moment t coneixem $AcVOL_t$, Pes_t i $Preu_t$, de manera que $IVOL_t$ es pot determinar fent servir la fórmula anterior.

En aquest punt sorgeix un problema clau. Els *swaps* es negocien en mercats extraborsaris i per aquest motiu, l'accés als preus és molt limitat. La fórmula anterior dóna una forma amb què es podrà treballar sense necessitat d'accedir directament als preus com veurem més endavant.

En estadis inicials del treball, l'objectiu era fer prediccions de S_t , a partir de dades d'anys anteriors, amb el mètode de Monte Carlo, tal com explica Glasserman (2014) [14]. Aquest és un mètode no determinista (és a dir, que té molts possibles resultats) que es fa servir per aproximar expressions complexes i costoses. Rep el seu nom en honor al Casino de Monte Carlo (Mònaco). La invenció i desenvolupament d'aquest mètode s'atribueix a Stanislaw Ulam i a John von Neumann, a la dècada dels anys quaranta.

En el nostre cas, usant el mètode de Black-Scholes, l'objectiu era fer N prediccions de S_t per a diferents t i diferents períodes de maduresa. Cada predicció rep el nom de *randomwalk* i pren una forma com la de la figura següent.

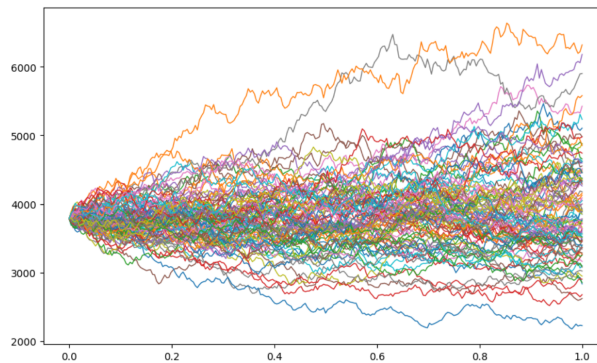


Figura 2: Exemple *randomwalks* amb el model Black-Scholes i actiu S: **S&P 500**

Amb aquests N preus obtinguts de la simulació de Monte Carlo, la idea és calcular la mitjana de les volatilitats (VOL_t) i amb aquesta, calcular $Preu_t = DF_t \times [\min(cap, VOL_t) - K]$ per aïllar $IVOL_t$ de l'equació (6.1).

Tanmateix, aquesta idea presentava el següent problema: donat que la volatilitat σ s'assumeix constant i coneguda en el model de Black-Scholes, per moltes simulacions de S_t que generés el model, la volatilitat mitjana que n'obtindríem seria més o menys constant i similar a la variable d'entrada σ introduïda. Per tant, el benefici d'aplicar aquest mètode no compensa el cost computacional que requereix.

Una alternativa més complexa seria fent servir un procés on la volatilitat no s'assumeix constant, com el model de Heston. Aquest model assumeix que el preu d'un actiu S_t ve determinat per un procés estocàstic

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dW_t^S,$$

on la volatilitat $\sqrt{\nu_t}$ segueix un procés anomenat procés d'Ornstein-Uhlenbeck

$$d\sqrt{\nu_t} = -\theta\sqrt{\nu_t}dt + \delta dW_t^\nu,$$

on W_t^S, W_t^ν són moviments Brownians.

Sense entrar gaire en detall mencionem que un procés d'Ornstein-Uhlenbeck és un procés de Gauss-Markov estacionari, és a dir, que satisfà les condicions d'un procés de Gauss i d'un procés de Markov.

Aquesta alternativa suposava un increment substancial de la complexitat dels models i, per tant, va ser descartada.

La solució proposada finalment va ser determinar tant K com $IVOL$ amb dades històriques, concretament amb dades compreses entre gener de 2015 i desembre de 2018, de la forma que es detalla a continuació.

6.1 Estimació de variables

Les estimacions que es proposen a continuació resulten exclusivament de les dades extretes d'internet sobre els actius **S&P 500** i **JPM** cotitzats a la borsa de Nova York.

6.1.1 Estimació de K

Com s'ha comentat en seccions prèvies, K es determina de forma que hi ha paritat a l'inici del contracte, és a dir, que $Preu_0 = 0$. Tenint en compte la fórmula (6.1), és immediat que $K = IVOL_0$. Aquesta variable no és res més que l'estimació o expectativa que té el mercat sobre la volatilitat realitzada que tindrà l'actiu subjacent a l'inici del contracte. Donades diverses madureses T (1, 3, 6, 9 i 12 mesos) definim K com el valor mitjà de les volatilitats realitzades de múltiples períodes de longitud T .

Aquests períodes es prenen amb superposició, és a dir, hi ha dades que serveixen com a *inputs* en diversos períodes. Més concretament, la superposició de cada període amb l'anterior és el següent:

- $T = 1$: superposició de $\frac{1}{2}$ del període (15 dies)
- $T = 3$: superposició de $\frac{2}{3}$ del període (2 mesos)
- $T = 6$: superposició de $\frac{2}{3}$ del període (4 mesos)
- $T = 9$: superposició de $\frac{2}{3}$ del període (6 mesos)
- $T = 12$: superposició de $\frac{2}{3}$ del període (8 mesos)

D'aquesta forma, per a cada T tindrem un preu K , determinat de forma històrica per a cada actiu subjacent. La figura 3 mostra gràficament les superposicions per a diverses madureses en el cas $t = 0$. Per determinar K es calcula la mitjana de les volatilitats realitzades de totes les marques vermelles.

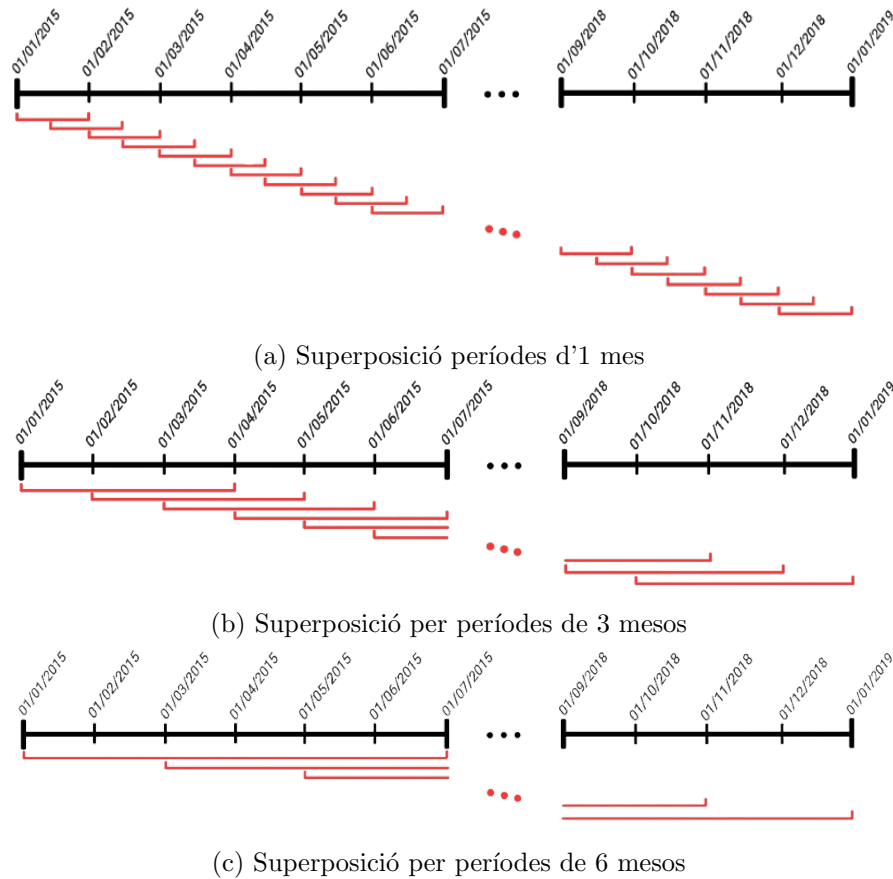


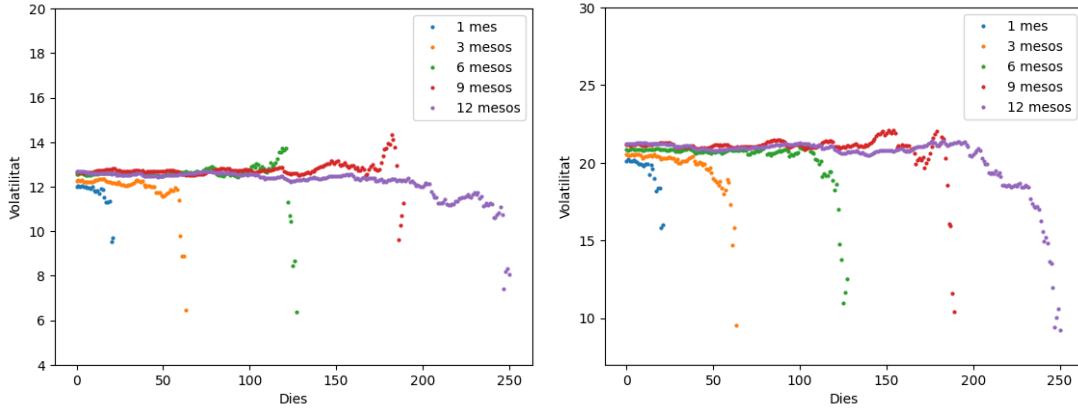
Figura 3: Superposicions per a diversos períodes

Els resultats obtinguts per l'actiu **S&P 500** són:

- Per $T = 1$ mes, $K = 12.002684$
- Per $T = 3$ mesos, $K = 12.277135$
- Per $T = 6$ mesos, $K = 12.558418$
- Per $T = 9$ mesos, $K = 12.648988$
- Per $T = 12$ mesos, $K = 12.693841$

6.1.2 Estimació d'IVOL

En cada moment $t > 0$, es té que el preu del contracte consta de dues parts, una corresponent a la volatilitat efectivament realitzada que correspon al període entre 0 i t , i una part no realitzada, corresponent al període comprès entre t i T . Per tal de seguir la metodologia utilitzada per determinar K , es determina $IVOL_t$ prenent el valor mitjà de les volatilitats realitzades de múltiples períodes de longitud $T - t$.



(a) IVOL per a diversos períodes del **S&P 500** (b) IVOL per a diversos períodes del **JPM**

Figura 4: IVOLs

La figura 4 mostra els resultats obtinguts per a les 5 madureses i els dos actius analitzats. S'observa clarament que per ambdós actius, la volatilitat sol variar poc al llarg del període fins que, al final dels contractes, la volatilitat fluctua bastant més i acaba en un descens molt pronunciat.

Aquest resultat és fàcilment explicable. És degut al fet que la longitud dels períodes no és igual per a tots, encara que la maduresa sigui la mateixa. Un període d'1 mes pot constar de 19, 20, 21, 22 i fins a 23 dies. Aleshores en fer el valor mitjà de les volatilitats realitzades dels dies 22 o 23 només s'hi introduiran tants valors com períodes de 22 o 23 dies que hi hagi, és a dir, la mitjana es calcula amb menys valors.

Això mateix passa amb les demès madureses i, per tant, els IVOL corresponents a t molt propers a T són calculats amb menys valors que els IVOL en t propers a 0. El que resulta curiós és que en tots els casos, a mesura que acaba el període, cap dels registres observats té una volatilitat molt elevada sinó que sempre és redueix fins a la meitat, aproximadament, dels valors inicials del període.

L'objectiu doncs és modelar IVOL en lloc de fer-ho directament amb *Preu* per veure fins a quin punt es pot precisar i predir la part del preu que és encara desconeguda.

7 Obtenció de variables d'entrada

En la primera part d'aquesta secció es consideraran algunes característiques distribucionals de l'actiu subjacent com la volatilitat, vista com la desviació estàndard, l'asimetria i la curtosi.

En la segona part s'introduirà la volatilitat implícita i, a partir d'aquesta, formes per construir l'asimetria seguint a Mixon (2011) [21] i la curtosi.

7.1 Moments implícits del mercat

Mitjançant el mètode dels moments es calcula la volatilitat esperada, l'asimetria i la curtosi.

Amb aquest mètode es pretén evitar l'ús de determinats models d'aproximació, com el que s'introduirà més endavant, que requereixen paràmetres o aproximacions inicials per començar a iterar. L'elecció d'aquesta aproximació inicial pot tenir una influència en el cost computacional del model així com la possibilitat que, per a diferents eleccions s'obtinguin diferents resultats finals.

Bakshi i Madan (2000) [2] i Carr i Madan (2001) [9] van demostrar que qualsevol funció de benefici (*payoff*) dues vegades diferenciable pot expressar-se de la forma:

$$f(S_T) = f(\kappa) + f'(\kappa)((S_T - \kappa)_+ - (\kappa - S_T)_+) + \int_0^\kappa f''(K)(K - S_T)_+ dK + \int_\kappa^\infty f''(K)(S_T - K)_+ dK$$

on κ és un `Strike` determinat. Prenent l'esperança sota la mesura neutral al risc de l'expressió anterior tenim

$$\mathbb{E}^Q[f(S_T)] = f(\kappa) + f'(\kappa)\exp(rT)(EC(\kappa, T) - EP(\kappa, T)) + \exp(rT)\left(\int_0^\kappa f''(K)EP(K, T)dK + \int_\kappa^\infty f''(K)EC(K, T)dK\right)$$

Ara, fent servir la paritat *put-call* (5.9), que estableix que $EC(K, T) - EP(K, T) = S_0 - K\exp(-rT)$ per a $t = 0$ podem escriure

$$\mathbb{E}^Q[f(S_T)] = f(\kappa) + f'(\kappa)\exp(rT)(S_0 - \kappa\exp(-rT)) + \exp(rT)\left(\int_0^\kappa f''(K)EP(K, T)dK + \int_\kappa^\infty f''(K)EC(K, T)dK\right)$$

Finalment, els resultats de Madan i Schoutens (2016) [20] permeten determinar una expressió tancada del moment m del retorn d'un actiu al llarg d'un període T , ja que prenent $f(S_T) = (\log(\frac{S_T}{S_0}))^m$ i el *forward price* $F_0 = S_0\exp(rT)$ tenim

$$\begin{aligned}
& \mathbb{E} \left[\log \left(\frac{S_T}{S_0} \right)^m \right] \\
&= \left(\log \left(\frac{\kappa}{S_0} \right) \right)^m + m \left(\log \left(\frac{\kappa}{S_0} \right) \right)^{m-1} \times \left(\frac{F_0}{\kappa} - 1 \right) \\
&+ \exp(rT) \int_0^\kappa \frac{m}{K^2} \times \left((m-1) \left(\log \left(\frac{K}{S_0} \right) \right)^{m-2} - \left(\log \left(\frac{K}{S_0} \right) \right)^{m-1} \right) \times EP(K, T) dK \\
&+ \exp(rT) \int_\kappa^\infty \frac{m}{K^2} \times \left((m-1) \left(\log \left(\frac{K}{S_0} \right) \right)^{m-2} - \left(\log \left(\frac{\kappa}{S_0} \right) \right)^{m-1} \right) \times EC(K, T) dK
\end{aligned}$$

amb taxa d'interès r , preu del forward F_0 , un Strike κ . $EC(K, T)$ i $EP(K, T)$ denoten el preu d'un *call Europeu* i d'una opció *put* respectivament de Strike K i maduresa T en anys. En la pràctica, els conjunts de preus Strike no són continus sinó que són conjunts discrets de valors, per tant podem aproximar la integral mitjançant el mètode dels trapezis. D'aquesta manera la fórmula anterior és equivalent a

$$\begin{aligned}
\mathbb{E} \left[\log \left(\frac{S_T}{S_0} \right)^m \right] &= \left(\log \left(\frac{K_0}{S_0} \right) \right)^m + m \left(\log \left(\frac{K_0}{S_0} \right) \right)^{m-1} \times \left(\frac{F_0}{K_0} - 1 \right) \\
&+ \exp(rT) m \sum_{i=1}^T \frac{\Delta K_i}{K_i^2} \left((m-1) \left(\log \left(\frac{K_i}{S_0} \right) \right)^{m-2} - \left(\log \left(\frac{K_i}{S_0} \right) \right)^{m-1} \right) \\
&\times Price(K_i, T)
\end{aligned}$$

on

- F_0 és el preu *forward* corrent, $F_0 = S_0 \exp(rT)$
- K_i és el i -èssim Strike *Out of The Money*: per a un *call* $K_i > F_0$ i per a un *put* $K_i < F_0$
- ΔK_i és l'interval entre Strikes:

$$\begin{cases} \Delta K_1 = K_2 - K_1 \\ \Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}, & \forall i \neq 1, m \\ \Delta K_m = K_m - K_{m-1} \end{cases} \quad (7.1)$$

- K_0 és el primer Strike per sota de F_0
- $Price(K_i, T)$ per l'opció amb Strike K_i i maduresa T ve donat per

$$\begin{cases} Price(K_i, T) = EP(K_i, T) & si \ K_i < K_0 \\ Price(K_i, T) = \frac{EC(K_0, T) + EP(K_0, T)}{2}, & si \ K_i = K_0 \\ Price(K_i, T) = EC(K_i, T) & si \ K_i > K_0 \end{cases} \quad (7.2)$$

on $EC(K, T)$ i $EP(K, T)$ es prenen tal que són el punt mitjà entre el preu *Bid* i el *Ask*.

Aquests valors *Bid* i *Ask* són molt utilitzats als mercats de valors i reflecteixen les dues parts principals que hi ha en un mercat, els compradors i els venedors. El *Bid* o preu de compra és el preu més alt que els inversors estan disposats a pagar, per tant, és el millor preu al què es pot vendre. De forma contrària, l'*Ask* o preu de venda fa referència al preu més baix al què els inversors estan disposats a vendre, en conseqüència, en aquest cas és el millor preu al què es pot comprar.

El preu de compra o *Bid* sempre és inferior al de venda o *Ask*. Això és degut al fet que quan es vol vendre, l'objectiu és fer-ho al preu més elevat possible i, quan es vol comprar, fer-ho al menor preu possible.

Sigui $R_T = \log(\frac{S_T}{S_0})$ i partint del resultat anterior, tenim que:

$$\begin{aligned} Var(R_T) &= \mathbb{E}[R_T^2] - (\mathbb{E}[R_T])^2 \\ Vol(R_T) &= \sqrt{Var(R_T)} = MI\ Vol. \\ Skew(R_T) &= \frac{\mathbb{E}[R_T^3] - 3\mathbb{E}[R_T]\mathbb{E}[R_T^2] + 2(\mathbb{E}[R_T])^3}{Var(R_T)^{3/2}} = MI\ Asim. \\ Kurt(R_T) &= \frac{\mathbb{E}[R_T^4] - 4\mathbb{E}[R_T]\mathbb{E}[R_T^3] + 6(\mathbb{E}[R_T])^2\mathbb{E}[R_T^2] - 3(\mathbb{E}[R_T])^4}{Var(R_T)^2} = MI\ Curt. \end{aligned}$$

7.1.1 Simulació d'inputs

Per tal d'aplicar la fórmula anterior, cal disposar d'observacions dels preus de *calls* i *puts* i els seus respectius preus *Strike*. La simulació d'aquests preus es fa amb les fórmules de Black-Scholes (5.7) i (5.8), introduïdes en la secció 4.

1. Es crea un conjunt $\kappa = \{K_i, i \in \mathbb{Z}\}$ format per valors enters on $K_i - K_{i-1} = l \in \mathbb{N}$ (l es determinarà en funció de l'actiu corresponent).
2. Per a tot T i t es determina $F_0 = S_0 e^{r(T-t)}$ i es determina K_0 l'element de κ que precedeix a F_0 .
3. Es limita el conjunt κ a una dimensió de $2n$ valors *Strike*, $\kappa = \{K_i, i = -n, \dots, n\}$ per sobre i per sota de K_0 .
4. Se simulen els preus de *calls* i *puts* mitjançant Black-Scholes tal que

$$\begin{aligned} EC_t(K_i, T) &= S_t \Phi(d_+) - K_i e^{-r(T-t)} \Phi(d_-) \quad \text{per } K_i > K_0 \\ EP_t(K_i, T) &= K_i e^{-r(T-t)} \Phi(-d_-) - S_t \Phi(-d_+) \quad \text{per } K_i < K_0 \end{aligned}$$

Finalment es reformula (7.1) de manera que

$$\begin{cases} \Delta K_{-n} = K_{-n+1} - K_{-n} \\ \Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}, & \forall i \neq -n, n \\ \Delta K_n = K_n - K_{n-1} \end{cases}$$

i aleshores

$$\begin{aligned}\mathbb{E}\left[\log\left(\frac{S_T}{S_0}\right)^m\right] &= \left(\log\left(\frac{K_0}{S_0}\right)\right)^m + m\left(\log\left(\frac{K_0}{S_0}\right)\right)^{m-1} \times \left(\frac{F_0}{K_0} - 1\right) \\ &\quad + \exp(rT)m \sum_{i=-n}^n \frac{\Delta K_i}{K_i^2} \left((m-1)\left(\log\left(\frac{K_i}{S_0}\right)\right)^{m-2} - \left(\log\left(\frac{K_i}{S_0}\right)\right)^{m-1} \right) \\ &\quad \times \text{Price}(K_i, T)\end{aligned}$$

Aplicant la fórmula proposada abans per a calcular $\mathbb{E}[R_T^m]$, per a cada $t = 0, \dots, T$ es calculen $\mathbb{E}[R_{T-t}]$, $\mathbb{E}[R_{T-t}^2]$, $\mathbb{E}[R_{T-t}^3]$, $\mathbb{E}[R_{T-t}^4]$ i amb aquests, *MI Vol.*, *MI Asim.* i *MI Curt.*

7.2 Volatilitat implícita

La volatilitat implícita és una mesura subjectiva i prospectiva (*ex-ante*) que reflecteix les expectatives que té el mercat sobre la volatilitat futura d'un actiu. A diferència de la volatilitat realitzada, aquesta se sol estimar mitjançant mètodes de valoració d'opcions, com el model de Black-Scholes. Usant aquest model, donat un moment t , el preu d'una opció *call* o *put* i els seus respectius *Strikes*, identifiquem un únic paràmetre lliure a la fórmula: el nivell de volatilitat.

Com que l'estructura de la fórmula no permet trobar una expressió tancada per trobar aquest valor, només a través d'aproximacions numèriques se'n pot trobar una solució. Un dels mètodes és el de Newton-Raphson. Aquest mètode ens proporciona la volatilitat depenent de la maduresa del contracte T i del *Strike*.

El mètode de Newton-Raphson és un algorisme per trobar aproximacions de zeros o arrels d'una funció real.

Sigui $f : [a, b] \rightarrow \mathbb{R}$ una funció derivable a l'interval real $[a, b]$. Sigui x_0 un valor inicial i $n \in \mathbb{N}$, aleshores es defineixen les iteracions del mètode per

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (7.3)$$

Per tal d'assegurar la convergència del mètode tenim el següent teorema:

Teorema 7.1. (Teorema de Convergència Local del Mètode de Newton)

Sigui $f \in \mathcal{C}^2([a, b])$. Si $p \in [a, b]$ i $f'(p) \neq 0$, aleshores existeix un $r > 0$ tal que el mètode de Newton-Raphson convergeix quadràticament a p per a tot $x_0 \in [p - r, p + r]$.

La demostració d'aquest teorema pot trobar-se en apunts d'assignatures com Mètodes Numèrics I.

La fórmula de Black-Scholes proporciona una eina de fixació de preus. L'objectiu és igualar aquesta fórmula (C_t) (5.7) amb les dades reals de preus d'opcions de què disposem (P_t), cal no confondre-ho amb la fórmula (5.8). Per tant, busquem les solucions σ tals que $P_t = C_t$ o el que és el mateix, arrels de la funció $0 = C_t - P_t$.

Sigui $f(\sigma) = C_t(\sigma) - P_t$. Aleshores s'obté el següent resultat:

Proposició 7.2. Sota el model de Black-Scholes, la fórmula d'iteració per a trobar la volatilitat implícita mitjançant el mètode de Newton-Raphson es defineix com

$$\sigma_{k+1} = \sigma_k - \frac{C_t(\sigma_k) - P_t}{S_t \Phi'(d_+) \sqrt{T-t}} \quad (7.4)$$

on S_t és el preu de l'actiu subjacent en el moment t , Φ la funció de distribució normal estàndard acumulada i P_t el preu del contracte en t .

Demostració. Donat que la fórmula de BS (Black-Scholes) $C_t(\sigma)$ ens proporciona el valor d'una opció de compra, volem que aquest sigui igual a P_t , per tant, busquem zeros de $f(\sigma) = C_t(\sigma) - P_t$. Per Newton-Raphson, tenim que

$$\sigma_{k+1} = \sigma_k - \frac{f(\sigma)}{f'(\sigma)}$$

El problema doncs es redueix a trobar $f'(\sigma)$ (derivada respecte σ).

$$\begin{aligned} f'(\sigma) &= \frac{\partial C(\sigma)}{\partial \sigma} = \frac{\partial}{\partial \sigma} (S_t \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-)) \\ &= S_t \frac{\partial \Phi}{\partial \sigma}(d_+) \frac{\partial d_+}{\partial \sigma} - K e^{-r(T-t)} \frac{\partial \Phi}{\partial \sigma}(d_-) \frac{\partial d_-}{\partial \sigma} \end{aligned}$$

Tenim que $d_{\pm} = \frac{\ln(\frac{S_t}{K}) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$ aleshores és fàcil veure que

$$\frac{\partial d_+}{\partial \sigma}(\sigma) = \frac{-\ln(\frac{S_t}{K}) - (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma^2\sqrt{T-t}} = -\frac{1}{\sigma}d_-$$

De forma anàloga obtenim que $\frac{\partial d_-}{\partial \sigma}(\sigma) = -\frac{1}{\sigma}d_+$. Per últim, com que $\Phi'(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ aleshores reescriuim la derivada com

$$\begin{aligned} f'(\sigma) &= S_t \frac{1}{\sqrt{2\pi}} e^{-d_+^2/2} \left(\frac{-d_-}{\sigma} \right) - K e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} e^{-d_-^2/2} \left(\frac{-d_+}{\sigma} \right) \\ &= \frac{1}{\sqrt{2\pi}} e^{-d_+^2/2} \left[-\frac{S_t d_-}{\sigma} + \frac{K e^{-r(T-t)} d_+}{\sigma} e^{(d_+^2 - d_-^2)/2} \right] \\ &= \frac{1}{\sqrt{2\pi}} e^{-d_+^2/2} \left[-\frac{S_t d_-}{\sigma} + \frac{K e^{-r(T-t)} d_+}{\sigma} \frac{S_t}{K} e^{-r(T-t)} \right] \\ &= \frac{1}{\sqrt{2\pi}} e^{-d_+^2/2} \left[-\frac{S_t d_-}{\sigma} + \frac{S_t d_+}{\sigma} \right] \\ &= S_t \Phi'(d_+) \left[\frac{d_+ - d_-}{\sigma} \right] = S_t \Phi'(d_+) \sqrt{T-t} \end{aligned}$$

La tercera igualtat s'obté del fet que $\frac{d_+^2 - d_-^2}{2} = \ln(\frac{S_t}{K}) + r(T-t)$ i l'última mitjançant que $d_- = d_+ - \sigma\sqrt{T-t}$ i per tant, $d_+ - d_- = d_+ - (d_+ - \sigma\sqrt{T-t}) = \sigma\sqrt{T-t}$. $\square$

El problema doncs és determinar el valor inicial σ_0 amb què començar les iteracions. Pel teorema anterior, tenim que per un σ_0 prou pròxim al valor buscat, el mètode convergirà. Per tant, considerarem expressions tancades d'estimació de la volatilitat implícita com la de Brenner & Subrahmanyam (1988) [7] o la de Corrado i Miller (1996) [10]. Aquestes fórmules aproximen bé el nivell de volatilitat quan el preu del subjacent S_t és similar a el Strike.

Brenner i Subrahmanyam proposen dues expressions senzilles en funció si el preu de l'actiu subjacent S és igual al preu descomptat del Strike K , això és (Ke^{-rT}) . En el primer cas tenim que:

$$\sigma_{BS} \approx \sqrt{\frac{2\pi}{T} \frac{C}{S}} \quad (7.5)$$

en el segon cas en què $S \neq Ke^{-rT}$ aleshores:

$$\sigma_{BS} \approx \sqrt{\frac{2\pi}{T} \frac{C - \delta}{S}} \quad (7.6)$$

on C és el preu de l'opció *call*, $\delta = \frac{S-X}{2}$ i $X = Ke^{-rT}$.

La proposta de Corrado i Miller és menys senzilla que l'anterior

$$\sigma_{CM} \approx \sqrt{\frac{2\pi}{T} \frac{1}{S-X} \left[C - \delta + \sqrt{(C - \delta)^2 - \frac{4\delta^2}{\pi}} \right]} \quad (7.7)$$

i donada la presència d'una arrel quadrada, és clar que la fórmula no està definida si $C \in (\delta - \frac{2\delta}{\sqrt{\pi}}, \delta + \frac{2\delta}{\sqrt{\pi}})$

El resultat obtingut en resoldre iterativament (7.4) és una superfície de tres eixos: maduresa T , Strike K i volatilitat implícita. Aquesta s'anomena *superfície de volatilitat*, i té un aspecte com el que s'observa a la figura 5a.

Fixant la variable T , obtenim, per a cada T , una corba que mostra la volatilitat implícita en funció de el Strike. Aquesta es coneix com la *volatility skew* o *volatility smile* segons la forma que prengui, com podem observar a la figura 5b.

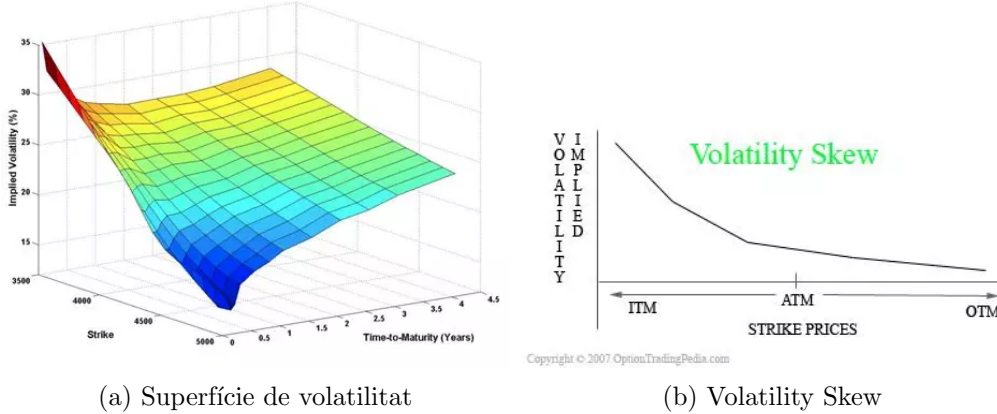


Figura 5: Figures de volatilitat

Escollirem com a *input 100% IV* el nivell de volatilitat implícita al 100% de *moneyness* (At The Money). Això significa que escollirem la volatilitat que situa el Strike al mateix nivell de S_t .

El nivell de volatilitat o *moneyness* és la posició relativa de l'actiu subjacent en el moment t , respecte del Strike del derivat. Aleshores les opcions es poden trobar en tres grups:

1. Out of The Money (OTM), quan $S_t < K$
2. At The Money (ATM), quan $S_t = K$
3. In The Money (ITM), quan $S_t > K$

És evident que al final del contracte s'obtenen beneficis positius si l'opció està ITM.

Ara considerarem només opcions *call* i suposarem fixat S_t .

Per exemple, si $S_t = K = 100$ aleshores l'opció és At The Money (*100% moneyness*), prendrem el nivell de volatilitat implícita corresponent com s'observa a la figura 6a.

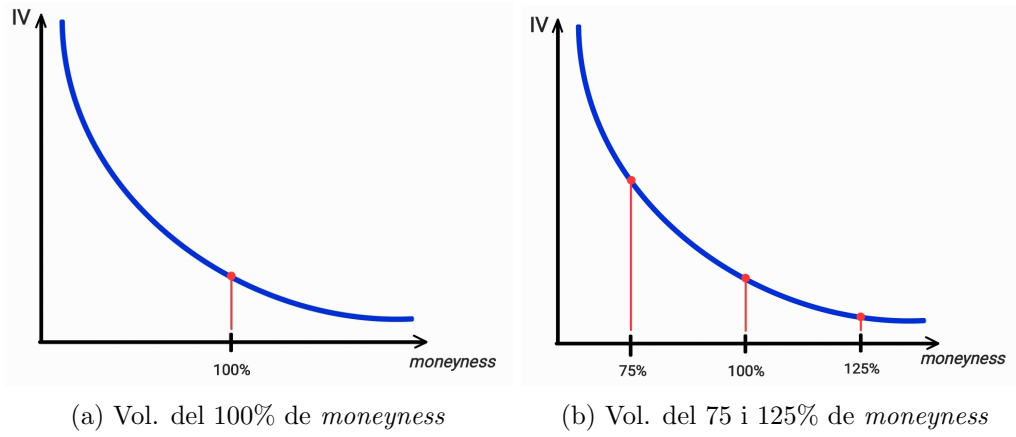


Figura 6: Nivells de volatilitat

Considerem ara formes d'incorporar mesures d'asimetria i curtosi. Per a l'asimetria, seguint a Mixon (2011) [21], una forma alternativa de construir-la és considerant la diferència entre el 75% i el 125% del nivell de volatilitat implícita respecte del 100%, l'anomenarem *IV Asim.*

Com es pot observar en la figura 6b, tenim que, per un S_t fix, en augmentar K , augmenta el nivell de *moneyness*, ja que ens acostem a nivells OTM (en context d'opcions *call*). De forma contrària, si K disminueix respecte S_t , aleshores el nivell de *moneyness* disminueix. Per convenció es determina que el numerador és la variable mòbil i el denominador la variable fixa. D'aquesta forma, podem calcular el 75% i el 125% buscant els K que satisfan

$$\frac{K}{S_t} = 0,75 \text{ i } \frac{K}{S_t} = 1,25, \quad (7.8)$$

Una vegada tenim els K , només cal determinar quina volatilitat els correspon a cada un. Denotem IV_{100} , IV_{75} i IV_{125} a les volatilitats implícites corresponents a cada nivell de *moneyness*.

Finalment, la diferència entre les volatilitats implícites del 75% i el 125% dividida entre la volatilitat implícita corresponent al 100%

$$IV_{Asim.} = \frac{IV_{75} - IV_{125}}{IV_{100}}$$

determina la mesura de l'asimetria proposada per Mixon (2011) [21].

De forma similar, i tot i que no hi ha una manera alternativa clara i definida de determinar la curtosi a partir del nivell de volatilitat implícita, en l'article de Höcht, Schoutens i Verschueren (2024) [16] es proposa examinar les cues de la corba per mesurar-la. Així doncs, la cua dreta i esquerra es defineixen com la diferència entre el 90 i el 95% respecte del 100% del nivell de volatilitat implícita i la diferència entre el 105 i el 110% respecte del 100% del nivell de volatilitat implícita respectivament. L'anomenarem *IV Curt*.

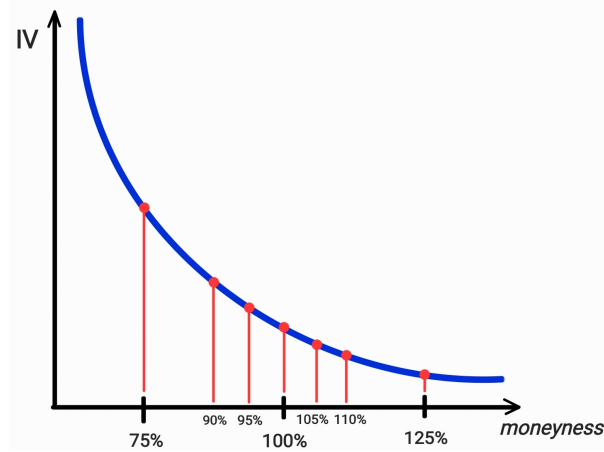


Figura 7: Vol. del 90, 95, 105 i 110% amb les prèviament descrites 75, 100 i 125%

7.2.1 Simulació d'inputs

Per tal d'aplicar el mètode de Newton-Raphson cal disposar de preus de *calls*. Aquest tipus de dades no té un accés tan restringit com els preus de *swaps* però com que la fórmula de Black-Scholes ens proporciona un resultat teòric per aquests preus la farem servir. Recordant la Proposició 7.2, tenim que cal igualar el preu dels *calls* amb la fórmula de Black-Scholes (5.7), per tant

$$S_t \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-) = P_t = C_t = S_t \Phi(d'_+) - K' e^{-r(T-t)} \Phi(d'_-) \quad (7.9)$$

on denotem d_+ , d_- i K als valors corresponents a la P_t i d'_+ , d'_- i K' als corresponents a C_t .

Si determinem aquests preus amb la mateixa fórmula, resulta que tenim la fórmula de Black-Scholes a ambdós costats de la igualtat.

Per tant, per poder aplicar Newton-Raphson cal determinar K en ambdós costats de la igualtat i determinar σ per al càlcul del preu del *call* (és a dir, al costat 'dret' de la igualtat), ja que l'objectiu és trobar la volatilitat (σ) corresponent a cada nivell de *moneyness* (m) (costat 'esquerre').

Per la fórmula (7.8) podem determinar K tal que $K = mS_t$, aleshores per a cada m tenim que

$$P_t = S_t \Phi(d_+) - mS_t e^{-r(T-t)} \Phi(d_-) = S_t \Phi(d'_+) - S_t e^{-r(T-t)} \Phi(d'_-) = C_t$$

i el K' corresponent a la banda dreta de la igualtat (7.9) el fixarem en $K' = S_t$ per a tot *moneyness* m .

Per últim, el preu K o **Strike**, reflecteix l'esperança o expectativa d'evolució d'un actiu, per tant, introduïrem aquesta característica en forma de capitalització per un període de temps igual al que hi ha entre el moment t i el final del contracte T , de forma que $K = mS_t e^{r(T-t)}$ per al càlcul de P_t i $K' = S_t e^{r(T-t)}$ per a l'altre costat de la igualtat. Per últim, es determina σ com la volatilitat històrica del període, en aquest cas, des del 2015 fins al 2018.

D'aquesta manera, (7.9) és

$$S_t \Phi(d_+) - mS_t \Phi(d_-) = S_t \Phi(d'_+) - S_t \Phi(d'_-)$$

i la funció de la qual cal trobar arrels amb Newton-Raphson és

$$\begin{aligned} f(\sigma) &= S_t \Phi(d'_+) - S_t \Phi(d'_-) - S_t \Phi(d_+) + mS_t \Phi(d_-) = \\ &= S_t (\Phi(d'_+) - \Phi(d_+)) + (m-1) S_t (\Phi(d_-) - \Phi(d'_-)) \\ &= \frac{S_t}{\sqrt{2\pi}} \int_{d_+}^{d'_+} e^{-z^2/2} dz + (m-1) \frac{S_t}{\sqrt{2\pi}} \int_{d'_-}^{d_-} e^{-z^2/2} dz \end{aligned}$$

Observació 7.3. Sota aquestes condicions, la volatilitat corresponent al 100% *moneyness*, això és, *100% IV*, és la volatilitat històrica constant per a tot t introduïda per a fer els càlculs.

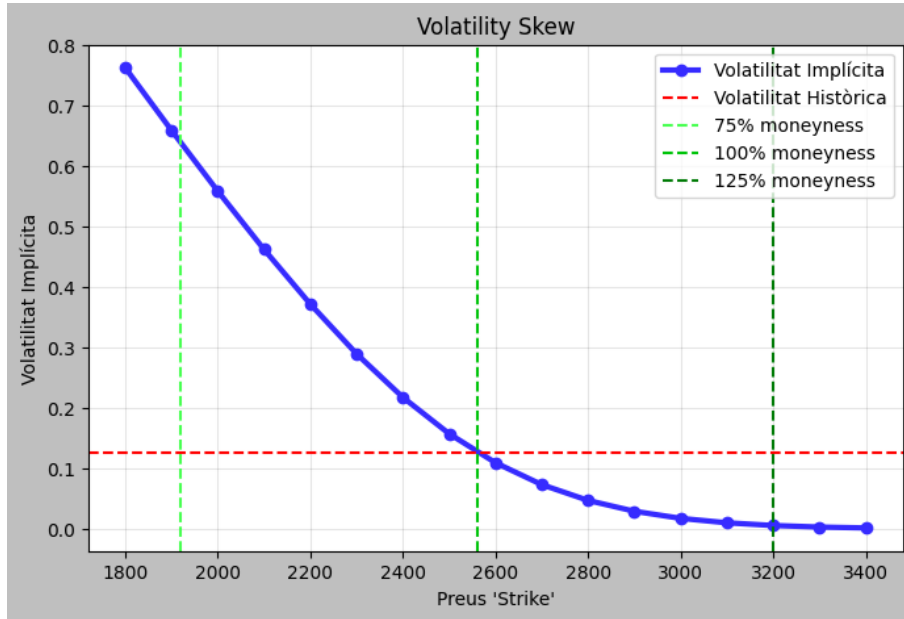


Figura 8: *Volatility skew* per a $T = 1$ any en $t = 0$ del **S&P 500** amb inici al 01.01.2015

Aquests valors serviran com a *inputs* per estimar *IVOL*. La justificació a la introducció d'aquestes mesures alternatives es troba en la correlació que hi ha entre ambdós mètodes, especialment en nivells baixos de maduresa T , com s'observa a la figura 9.

	MI Vol & 100% IV	MI Skew & (75-125/100)% IV	MI Kurt & (90-95/100)% IV	MI Kurt & (105-110/100)% IV
1 mes	-0.869290	-0.978719	0.943622	-0.554539
3 mesos	-0.846302	-0.969599	0.870196	-0.434293
6 mesos	-0.839354	-0.949436	0.807855	-0.465357
9 mesos	-0.835709	-0.911974	0.606349	-0.279852
12 mesos	-0.835750	-0.891867	0.636898	-0.286897

Figura 9: Correlació entre variables d'ambdós mètodes

Per últim, agrupem les dades en diferents conjunts.

S1 conté l'estimació de la volatilitat fent servir el mètode dels moments (**MI Vol.**),
S2 conté **MI Vol.** i la mesura d'asimetria (**MI Asim.**),
S3 conté **MI Vol.**, **MI Asim.** i la mesura de curtosis (**MI Curt.**).

S4 conté els nivells de volatilitat del 100% de *moneyness* (**100% IV**),
S5 conté **100% IV** i la mesura d'asimetria proposada per Mixon (**IV Asim.**) i
S6 conté **100% IV**, **IV Asim.** i la mesura de curtosis (**IV Curt.**).

(El terme IV es refereix a *Implied Volatility* i MI a *Market Implied*).

En resum, tenim les dades agrupades en 6 grups, els tres primers provenen del mètode del moments i els tres últims es basen en la volatilitat implícita del mercat.

S1: MI Vol., **S2:** MI Vol. & MI Asim., **S3:** MI Vol. & MI Asim. & MI Curt.
S4: 100% IV, **S5:** 100% IV & IV Asim., **S6:** 100% IV & IV Asim. & IV Curt.

8 Models predictius

En aquesta secció es proposen diferents models per fer les estimacions. Considerarem la Regressió del Procés Gaussià, tècniques d'aprenentatge automàtic basades en arbres i la Regressió Lineal.

8.1 Regressió del Procés Gaussià

En aquesta secció farem una introducció a la Regressió del Procés Gaussià seguint Ras-
munssen i Williams (2006) [22].

Definició 8.1. *Un procés de Gauss o Gaussià és una col·lecció de variables aleatòries de les quals qualsevol subcol·lecció finita té una distribució Gaussiana (normal) conjunta.*

Sigui $\mathcal{D} = (\mathbf{X}, \mathbf{y}) = \{(\mathbf{x}_i, y_i) \mid i = 1, \dots, n\}$ un grup d'observacions on $\mathbf{x}_i$ és un vector d'entrada de dimensió p i y_i el corresponent valor escalar de sortida. Per tot $i = 1, \dots, n$

$$y_i = f(\mathbf{x}_i) + \epsilon_i, \quad \text{on } \epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$$

Si assumim que f és un procés de Gauss aleshores $f \sim \mathcal{GP}(m(x), k(x, x))$. El procés està completament caracteritzat per la funció mitjana m , sovint definida com la funció zero, i la funció covariància o kernel k . A més, les observacions també seran un procés Gaussià $\mathbf{y} \sim \mathcal{GP}(0, K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 Id)$ on

$$K(\mathbf{X}, \mathbf{X}) = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k(\mathbf{x}_2, \mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & k(\mathbf{x}_n, \mathbf{x}_2) & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix}$$

Per un nou grup d'observacions $\mathbf{X}^*$, de dimensió $n^* \times p$ i els corresponents valors desconeguts f^* , aleshores la probabilitat conjunta segueix una distribució normal multivariant,

$$\begin{bmatrix} \mathbf{y} \\ f^* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(\mathbf{X}, \mathbf{X} + \sigma_n^2 Id) & K(\mathbf{X}, \mathbf{X}^*) \\ K(\mathbf{X}^*, \mathbf{X}) & K(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix}\right) \dim (n + n^*) \times (n + n^*).$$

ja que $K(\mathbf{X}, \mathbf{X}^*)$ té dimensió $n \times n^*$, $K(\mathbf{X}^*, \mathbf{X})$ té dimensió $n^* \times n$ i $K(\mathbf{X}^*, \mathbf{X}^*)$ dimensió $n^* \times n^*$.

Sigui $Y = (Y_1, Y_2, \dots, Y_n)^T$ una variable aleatòria amb funció de densitat

$$f_Y(y) = \frac{1}{(\sqrt{2\pi})^n |\det \Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (y - b)^T \Sigma^{-1} (y - b) \right\}$$

on $y = (y_1, y_2, \dots, y_n)^T$, $b \in \mathbb{R}$ i Σ és una matriu definida positiva de dimensió $n \times n$. Aleshores diem que Y és una distribució normal n -dimensional amb mitjana b i matriu de covariàncies Σ . Escrivim $Y \sim \mathcal{N}_n(b, \Sigma)$.

Proposició 8.2. *Sigui $Y = \begin{pmatrix} Y_{(1)} \\ Y_{(2)} \end{pmatrix} \begin{matrix} \} m \\ \} n - m \end{matrix}$ si $\text{Cov}(Y_{(1)}, Y_{(2)}) = 0$, aleshores $Y_{(1)}$ i $Y_{(2)}$ són independents i amb distribució normal.*

Demostració. Se segueix de la factorització de la *pdf*. □

Proposició 8.3. *Sigui $Y = \begin{pmatrix} Y_{(1)} \\ Y_{(2)} \end{pmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix}$, $Y \sim \mathcal{N}_n(b, \Sigma)$ amb*

$$b = \begin{pmatrix} b_{(1)} \\ b_{(2)} \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix}$$

Aleshores

$$Z = Y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}$$

és independent de $Y_{(1)}$.

Demostració. És suficient veure que $\text{Cov}(Y_{(1)}, Z) = 0$. En particular

$$\begin{aligned} \text{Cov}(Y_{(1)}, Z) &= \text{Cov}(Y_{(1)}, Y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}) \\ &= \text{Cov}(Y_{(1)}, Y_{(2)}) - \text{Cov}(Y_{(1)}, Y_{(1)})\Sigma_{11}^{-1}\Sigma_{12} \\ &= \Sigma_{12} - \Sigma_{11}\Sigma_{11}^{-1}\Sigma_{12} = 0 \end{aligned}$$

Ara només cal aplicar la proposició anterior. □

Proposició 8.4. *Amb la mateixa notació que abans,*

$$Y_{(2)}|Y_{(1)} = y_{(1)} \sim \mathcal{N}_{n-m}(b_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(y_{(1)} - b_{(1)}), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

Demostració. Per la proposició anterior, tenim que

$$f_{(Y_{(1)}, Z)}(y_{(1)}, z) = f_{Y_{(1)}}(y_{(1)})f_Z(z),$$

a més, fent el canvi $(Y_{(1)}, Y_{(2)}) \rightarrow (Y_{(1)}, Z)$, es té

$$f_{(Y_{(1)}, Z)}(y_{(1)}, z) = f_{(Y_{(1)}, Y_{(2)})}(y_{(1)}, z + \Sigma_{21}\Sigma_{11}^{-1}y_{(1)}),$$

i, per tant,

$$f_Z(z) = \frac{f_{(Y_{(1)}, Y_{(2)})}(y_{(1)}, z + \Sigma_{21}\Sigma_{11}^{-1}y_{(1)})}{f_{Y_{(1)}}(y_{(1)})}$$

això és

$$\begin{aligned} f_Z(y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}y_{(1)}) &= \frac{f_{(Y_{(1)}, Y_{(2)})}(y_{(1)}, y_{(2)})}{f_{Y_{(1)}}(y_{(1)})} \\ &= f_{Y_{(2)}|Y_{(1)}}(y_{(2)}|y_{(1)}) \end{aligned}$$

Ara,

$$Z \sim \mathcal{N}_{n-m}(b_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}b_{(1)}, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) \quad (8.1)$$

En particular,

$$\mathbb{E}(Z) = \mathbb{E}(Y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}) = \mathbb{E}(Y_{(2)}) - \Sigma_{21}\Sigma_{11}^{-1}\mathbb{E}(Y_{(1)})$$

i també

$$\begin{aligned} \text{Cov}(Z, Z) &= \text{Cov}(Y_{(2)}, Y_{(2)}) - \text{Cov}(Y_{(2)}, \Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}) \\ &\quad - \text{Cov}(\Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}, Y_{(2)}) + \text{Cov}(\Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}, \Sigma_{21}\Sigma_{11}^{-1}Y_{(1)}) \\ &= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{11}\Sigma_{11}^{-1}\Sigma_{12} \\ &= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{aligned}$$

de tal manera que $f_Z(y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}y_{(1)})$ és la funció de densitat de

$$\mathcal{N}_{n-m}(b_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(y_{(1)} - b_{(1)}), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

a més, tenint en compte (8.1) i posant $\bar{\Sigma} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ tenim

$$\begin{aligned} f_Z(y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}y_{(1)}) &= \\ &= \frac{1}{(\sqrt{2\pi})^n |\det \bar{\Sigma}|^{1/2}} \\ &\times \exp\left\{-\frac{1}{2}(y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}y_{(1)} - (b_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}b_{(1)}))^T \bar{\Sigma}^{-1}(y_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}y_{(1)} - (b_{(2)} - \Sigma_{21}\Sigma_{11}^{-1}b_{(1)}))\right\} \\ &= \frac{1}{(\sqrt{2\pi})^n |\det \bar{\Sigma}|^{1/2}} \\ &\times \exp\left\{-\frac{1}{2}(y_{(2)} - (b_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(y_{(1)} - b_{(1)}))^T \bar{\Sigma}^{-1}(y_{(2)} - (b_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(y_{(1)} - b_{(1)})))\right\} \end{aligned}$$

□

Usant la Proposició obtenim que la distribució condicionada de f^* donats $\mathbf{X}$, $\mathbf{y}$ i $\mathbf{X}^*$ és

$$\begin{aligned} f^*|\mathbf{X}^*, \mathbf{X}, \mathbf{y} &\sim \mathcal{N}(K(\mathbf{X}^*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 Id]^{-1}\mathbf{y}, \\ &K(\mathbf{X}^*, \mathbf{X}^*) - K(\mathbf{X}^*, \mathbf{X}) \times [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 Id]^{-1}K(\mathbf{X}, \mathbf{X}^*)) \end{aligned}$$

Aquest resultat dona una forma de calcular f^* , ja que la funció mitjana de la fórmula anterior serà la que fem servir per a l'estimació de f^* . Aquesta funció depèn de l'elecció de la funció de covariància k . Veiem ara quines funcions k existeixen.

Una opció molt comuna és el *squared exponential* kernel $k: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$, donat per

$$k(\mathbf{x}, \mathbf{x}^*) = \sigma^2 \exp\left(-\frac{|\mathbf{x} - \mathbf{x}^*|^2}{2l^2}\right),$$

amb el conjunt de paràmetres $\boldsymbol{\theta} = (\sigma^2, l)$. Com a alternativa, també es pot considerar el Matern kernel, que és una generalització de l'anterior,

$$k(\mathbf{x}, \mathbf{x}^*) = \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{l}|\mathbf{x} - \mathbf{x}^*|\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}}{l}|\mathbf{x} - \mathbf{x}^*|\right),$$

on K_ν és una funció de Bessel modificada i $\boldsymbol{\theta} = (\nu, l)$.

Les funcions de covariància o kernels proporcionen la correlació existent entre les observacions. El Matern kernel, en particular, permet ajustar el grau de suavitat entre les observacions amb el paràmetre ν . Quan aquest paràmetre és 1/2 aleshores el Matern kernel es comporta com el *squared exponential* kernel. En concret, com més elevat és ν , més suau és el model i com més petit, més rugós, com es pot apreciar a la figura 10. El paràmetre l controla la distància en què dues observacions poden considerar-se correlacionades. Un valor de l baix implica que només observacions properes estan relacionades i un valor de l alt permet correlacions entre observacions més allunyades.

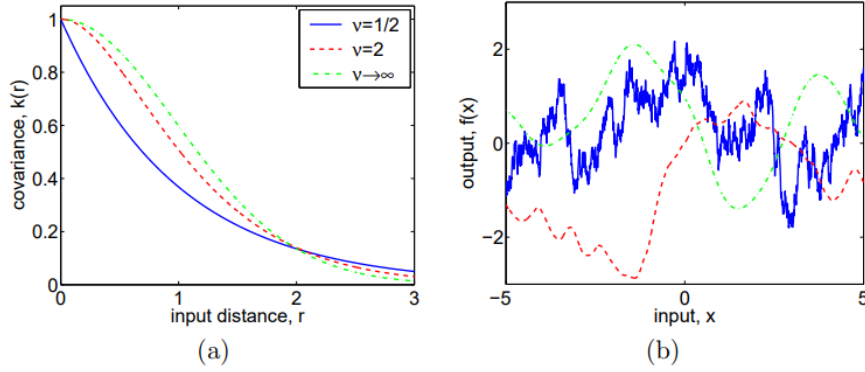


Figura 10: (a) funcions de covariància en funció de la distància i (b) el resultat del procés de Gauss amb el Matern kernel, per diferents valor de ν , amb $l = 1$

La funció de Bessel que es fa servir en el Matern kernel s'anomena funció de Bessel modificada de segona espècie d'ordre ν , K_ν . Aquesta funció permet ajustar el nivell de correlació entre punts en funció de la seva distància. Té un decreixement exponencial que regula la disminució de la correlació entre observacions a mesura que augmenta distància entre elles. Aquestes funcions són molt utilitzades per modelar dades temporals, ja que permeten ajustar la dependència de distància de forma flexible.

Les funcions de Bessel són solucions canòniques $y(x)$ de l'equació diferencial de Bessel:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - \nu^2)y = 0$$

El nombre ν s'anomena l'ordre de les funcions de Bessel associades a l'equació. Tanmateix, les funcions de Bessel modificades I_ν i K_ν són solucions de l'equació de Bessel modificada

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (x^2 + \nu^2)y = 0$$

Les funcions de Bessel modificades de primera espècie ($I_\nu(x)$) d'ordre ν venen donades per

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(k + \nu + 1)} \left(\frac{x}{2}\right)^{2k + \nu}$$

i les de segons espècie ($K_\nu(x)$) d'ordre ν venen donades per

$$K_\nu(x) = \frac{\pi}{2} \frac{I_{-\nu}(x) - I_\nu(x)}{\sin(\nu\pi)}$$

Aquestes funcions són exponencialment creixents i decreixents respectivament (vegeu figura 11), és per això que prenem K_ν per tal de regular un decreixement de la correlació entre observacions a mesura que aquestes són llunyanes entre elles.

Per últim, cal determinar quins valors pren el conjunt θ . Una forma comuna d'estimar-lo és maximitzant el *marginal log-likelihood* de la distribució Gaussiana multivariant sobre el conjunt d'entrenament $(\mathbf{X}, \mathbf{y})$

$$\log L(\mathbf{y}|\mathbf{X}, \theta) = -\frac{1}{2} \mathbf{y}^T K(\mathbf{X}, \mathbf{X})^{-1} \mathbf{y} - \frac{1}{2} \log[\det(K(\mathbf{X}, \mathbf{X}))] - \frac{n}{2} \log(2\pi)$$

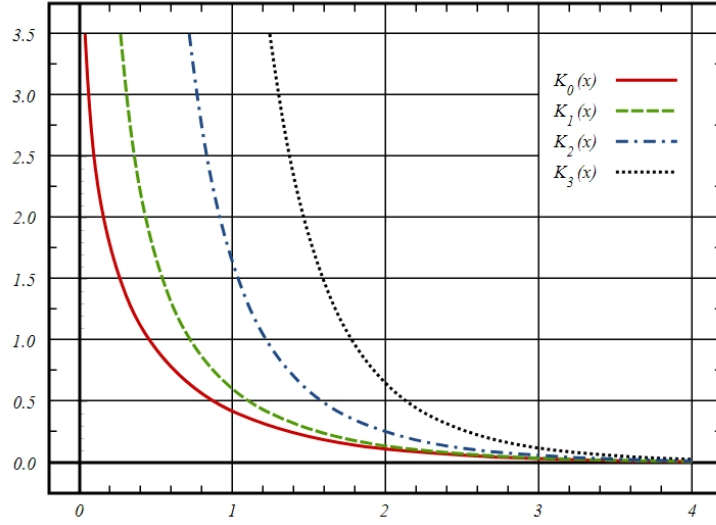


Figura 11: Funció de Bessel modificada de segona espècie per a diferents ν , $K_\nu(x)$

8.2 Aprenentatge automàtic basat en arbres

L'aprenentatge automàtic basat en arbres és una branca de l'aprenentatge automàtic que utilitza estructures d'arbres per a prendre decisions i fer prediccions. Els arbres de decisió, introduïts per Breiman, Friedman, Stone i Olshen (1984) [6], parteixen l'espai de predicció en regions disjunctes de tal forma que, per a cada element de cada subregió $\{R_j \mid j = 1, \dots, J\}$ es prediu el mateix valor $\hat{y}_{R_j}$;

$$f_{arbre}(\mathbf{x}) = \sum_{j=1}^J \hat{y}_{R_j} \cdot \mathbb{1}_{(\mathbf{x} \in R_j)} \quad (8.2)$$

on $\hat{y}_{R_j}$ és la mitjana de les observacions d'entrenament en la regió R_j . Conceptualment, les fulles dels arbres representen les prediccions.

Per un espai de característiques de dimensió p , un subespai de $\mathbb{R}^p$, l'espai es divideix en J regions, R_j sent regions o hiperblocs p -dimensionals.

Aquests arbres es construeixen mitjançant l'anomenat algorisme de CART (*Classification and Regression Trees*). Els punts frontera o *split points* es determinen analitzant, per cada observació, quina és la que té un valor més petit de la funció pèrdua $\mathcal{L}$.

$$\mathcal{L}(y, f(\mathbf{x})) = \frac{1}{n} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2. \quad (8.3)$$

Aleshores, de totes les característiques i observacions possibles, prenem la que té una funció pèrdua més petita. Aquesta característica, junt amb el valor de l'observació, serà l'arrel (*root*) de l'arbre de regressió. Un cop tenim la primera separació tenim dues regions. Podem doncs repetir el procés de nou dins de cada nova regió iterativament. Així, obtindrem un arbre amb el mateix nombre de fulles que d'observacions i, tot i que tindran un biaix baix, la variància serà elevada i la capacitat predictiva per a noves observacions serà més aviat dolenta.

És per això que s'introdueixen criteris de parada, per exemple quan una regió té un nombre d'observacions inferior a un nombre m determinat.

Aquest algorisme es defineix com a *greedy* o voraç. Això significa que en cada iteració, l'algorisme escull la millor opció per a cada pas amb la informació de la qual disposa en aquell moment, en comptes de 'mirar endavant' per determinar una solució global que podria resultar millor. L'algorisme no es replanteja en el futur les decisions preses i és per això que el resultat podria no ser òptim.

Els arbres de regressió per si mateixos estan lluny de ser bons mètodes predictors, a causa de la sensibilitat a canvis de dades, a la propensió al *overfitting* o sobreajust (es fan massa regions) i a l'elevada correlació entre les prediccions. Per tal de millorar la capacitat predictiva, els arbres de predicció s'agrupen en conjunts d'arbres.

Considerarem dos mètodes: el bosc aleatori (*random forest*) i la potenciació del gradient (*gradient boosting machine*).

Aquests algorismes presenten avantatges com són la fàcil comprensió i interpretabilitat així com una menor sensibilitat a valors extrems, com és el cas d'altres models com la regressió. Tanmateix, les limitacions principals, com en el cas dels arbres de regressió per si sols, són la tendència al sobreajust i la baixa estabilitat, que fa que canvis petits en les dades puguin generar arbres completament diferents.

8.3 Bosc Aleatori

Definició 8.5. *Un bosc aleatori és una combinació d'arbres predictors o de regressió tal que cada arbre depèn del valor d'un vector aleatori provat independentment i amb la mateixa distribució per a cada un d'aquests (arbres).*

Aquesta tècnica d'aprenentatge automàtic basada en arbres de decisió va ser introduïda per Leo Breiman l'any 2001 [5].

8.3.1 Bagging

Aquest bosc es genera mitjançant el que s'anomena *bagging* o *bootstrap aggregating technique*. Aquesta tècnica consisteix en, a partir d'un conjunt de dades inicials $\mathcal{D} = (\mathbf{X}, y) = \{(\mathbf{x}_i, y_i) \mid i = 1, \dots, n\}$, generar múltiples subconjunts aleatòriament $\mathcal{D}_t^b$ de mida $\delta \cdot n$, $0 < \delta < 1$. Aquests subconjunts generats aleatòriament segueixen la mateixa distribució subjacent que el conjunt $\mathcal{D}$.

Error quadràtic mig (EQM) o *mean squared error (MSE)*

L'error quadràtic mig d'un estimador o predictor mesura la mitjana dels errors al quadrat, és a dir, la diferència entre l'estimador $\hat{\theta}$ i el valor que s'estima θ . Mesura la qualitat de l'estimador i com més pròxim a zero sigui l'EQM millor serà l'estimador. A més, incorpora tant el biaix com la variància de l'estimador.

$$\begin{aligned} EQM &= \mathbb{E}((\hat{\theta} - \theta)^2) = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] + \mathbb{E}\left[(\mathbb{E}(\hat{\theta}) - \theta)^2\right] \\ &= Var(\hat{\theta}) + Bias(\hat{\theta}, \theta)^2 \end{aligned}$$

Demostració. És suficient veure que $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ i substituir X per $\hat{\theta} - \theta$.

$$\begin{aligned} Var(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned}$$

Per tant, $\mathbb{E}((\hat{\theta} - \theta)^2) = Var(\hat{\theta} - \theta) + (\mathbb{E}(\hat{\theta} - \theta))^2$. □

Gràcies a aquesta tècnica es redueix la variància del model comparada amb els arbres de regressió individualment degut al següent resultat de Breiman (1996) [4].

Sigui $f(x, \mathcal{D})$ un (arbre) predictor per a y , aleshores denotem $\{f(x, \mathcal{D}_t)\}$ la seqüència de predictors per a cada conjunt $\mathcal{D}_t$.

Sigui $(x, y) \in \mathcal{D}$ i sigui $f(x, \mathcal{D})$ el predictor de y , aleshores la predicció agregada és el valor esperat sobre $\mathcal{D}$ de $f(x, \mathcal{D})$, això és

$$\phi_A(x) = \mathbb{E}_{\mathcal{D}} f(x, \mathcal{D}) \quad (8.4)$$

Prenem x un valor *input* fix i y el valor de sortida corresponent. Aleshores

$$\mathbb{E}_{\mathcal{D}}(y - f(x, \mathcal{D}))^2 = y^2 - 2y\mathbb{E}_{\mathcal{D}} f(x, \mathcal{D}) + \mathbb{E}_{\mathcal{D}} f^2(x, \mathcal{D}) \quad (8.5)$$

fent servir (8.4) i aplicant la desigualtat $\mathbb{E}[Z^2] \geq (\mathbb{E}[Z])^2$ al tercer terme de (8.5) obtenim que

$$\mathbb{E}_{\mathcal{D}}(y - f(x, \mathcal{D}))^2 \geq (y - \phi_A(x))^2$$

Integrant en ambdós costats sobre la distribució conjunta de x, y obtenim que MSE de $\phi_A(x)$ és més baix que el MSE esperat sobre $\mathcal{D}$ de $f(x, \mathcal{D})$. Per tant tenim que per a qualsevol predictor $f(x, \mathcal{D})$, tindrem que el MSE serà major que per al predictor $\phi_A(x)$.

8.3.2 CART

L'algorisme CART vist en l'apartat anterior calcula per cada característica i cada observació la funció pèrdua $\mathcal{L}$ (8.3) per determinar quins seran els *split points* i les característiques arrel. Veiem com l'algorisme fa servir totes les característiques disponibles.

Els boscos aleatoris, introduïts per Breiman al 2001 [5], incorporen un altre element d'aleatorietat per tal de decidir quin serà el criteri de partició en cada node. En comptes de fer servir totes les característiques disponibles (inicialment p), es fan servir un nombre m de característiques ($m < p$) escollides aleatòriament. La introducció de l'aleatorietat fa que els arbres tinguin menys correlació entre ells.

8.3.3 Algorisme

1. A partir de $\mathcal{D} = (\mathbf{X}, y) = \{(\mathbf{x}_i, y_i) \mid i = 1, \dots, n\}$ es generen subconjunts aleatoris $\mathcal{D}_t^b = \{(\mathbf{x}_i^b, y_i^b) \mid i = 1, \dots, n\}$ on $t = 1, \dots, T$.
2. Per a cada subconjunt, es construeix l'arbre $f_b(x|\mathcal{D}_t^b)$ minimitzant la funció pèrdua $\mathcal{L}$ i utilitzant m característiques aleatòries de les p disponibles.

3. Finalment, la predicció s'obtéindrà prenent la mitjana dels resultats obtinguts en cada arbre individualment:

$$f_{rf}(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^T f_b(\mathbf{x} \mid \mathcal{D}_t^b)$$

8.4 Potenciació del Gradient (*gradient boosting*)

A diferència del mètode anterior que es basa en la construcció d'arbres independents, la tècnica de la potenciació del gradient consisteix en la creació de forma iterativa d'arbres que es basen en 'l'error' de l'anterior, Friedman (2001) [13].

En cada iteració t es construeix un arbre de dimensió limitada (*depth* o max fulles), aquest no es construeix directament sobre el valor y_i que volem predir, sinó usant el que anomenem els pseudo-residuals $\rho_{i,t}$ per l'observació i ,

$$\rho_{i,t} = - \left[\frac{\partial \mathcal{L}(y_i, f(\mathbf{x}_i))}{\partial f(\mathbf{x}_i)} \right]_{f=f_{t-1}}$$

L'elecció de la funció pèrdua és arbitrària però la més usada és $\mathcal{L}(y, f(x)) = \frac{1}{2}(y - f(x))^2$ per la senzillesa en derivar-la. D'aquesta forma, els pseudo-residus queden

$$\rho_{i,t} = [y_i - f(\mathbf{x}_i)]_{f=f_{t-1}}$$

Amb aquests valors $\rho_{i,t}$ es crea un arbre de regressió amb l'algorisme CART mencionat anteriorment. Donat que per a cada regió terminal creada $\{R_{j,t} \mid j = 1, \dots, J_t\}$, on J_t és el nombre de fulles de l'arbre en la iteració t , podem tenir més d'un pseudo-residu, es calcula el següent valor (per tal que cada fulla tingui un valor únic)

$$\hat{b}_{j,t} = \underset{b}{\operatorname{argmin}} \sum_{i: \mathbf{x}_i \in R_{j,t}} \mathcal{L}(y_i, f_{t-1}(\mathbf{x}_i) + b)$$

on es busca el valor b que minimitza l'expressió anterior. Aquest $b_{j,t}$ és el que determina quin valor té cada regió $R_{j,t}$ i, amb aquest, actualitzem la funció f_t de la manera següent

$$f_t(\mathbf{x}_i) = f_{t-1}(\mathbf{x}_i) + \lambda \sum_{j=1}^{J_t} \hat{b}_{j,t} \cdot \mathbb{1}_{(\mathbf{x}_i \in R_{j,t})} = f_{t-1}(\mathbf{x}_i) + \lambda f_{\text{arbre}_t}(\mathbf{x}_i)$$

El paràmetre λ controla la velocitat de l'algorisme i el procés no s'atura fins que t assoleix el nombre determinat d'arbres T o fins que la reducció de l'error deixa de ser significant en cada iteració nova. La predicció final de l'algorisme correspon amb $f_T(\mathbf{x})$ obtingut amb l'expressió anterior.

Per començar el mètode es parteix de $f_0(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n y_i$. Tenim que $f_0(\mathbf{x})$ es defineix com

$$f_0(\mathbf{x}) = \underset{b}{\operatorname{argmin}} \sum_{i=1}^n \mathcal{L}(y_i, b) = \underset{b}{\operatorname{argmin}} \left(\sum_{i=1}^n \frac{1}{2} (y_i - b)^2 \right)$$

Derivant l'expressió entre parèntesis respecte b i igualant-la a 0 per trobar el b mínim tenim que

$$0 = - \sum_{i=1}^n (y_i - b) = - \sum_{i=1}^n y_i + nb$$

I, per tant, obtenim el resultat que volíem.

De manera informal, podríem dir que en cada iteració creem un arbre amb els pseudo-residus, i actualitzem f_t en cada pas sumant els $t - 1$ arbres anteriors així com la mitjana inicial f_0 .

8.4.1 Algorisme

1. Partim de $f_0(x) = \operatorname{argmin}_b \sum_{i=1}^n \mathcal{L}(y_i, b)$. Com s'acaba de veure es té que $b = \frac{1}{n} \sum_{i=1}^n y_i$.
2. Per a $t = 1$ fins a T repetim el següent:
 - (a) Calculem els pseudo-residus per a t ,

$$\rho_{i,t} = - \left[\frac{\partial \mathcal{L}(y_i, f(\mathbf{x}_i))}{\partial f(\mathbf{x}_i)} \right]_{f=f_{t-1}} = [y_i - f(\mathbf{x}_i)]_{f=f_{t-1}}$$

- (b) Creem un arbre de regressió amb els pseudo-residus mitjançant l'algorisme CART descrit anteriorment (sense *bagging*).
Aleshores tenim les regions $\{R_{j,t} \mid j = 1, \dots, J_t\}$, on J_t és el nombre de fulles de l'arbre i els valors d'aquestes fulles són els pseudo-residus, on cada x_i pertany a una regió. El valor d'aquestes regions es determina de la següent forma.
 - (c) Determinem $\hat{b}_{j,t} = \operatorname{argmin}_b \sum_{i: \mathbf{x}_i \in R_{j,t}} \mathcal{L}(y_i, f_{t-1}(\mathbf{x}_i) + b)$
 - (d) Actualitzem f_t ,

$$f_t(\mathbf{x}_i) = f_{t-1}(\mathbf{x}_i) + \lambda \sum_{j=1}^{J_t} \hat{b}_{j,t} \cdot \mathbb{1}_{(\mathbf{x}_i \in R_{j,t})} = f_{t-1}(\mathbf{x}_i) + \lambda f_{\text{arbre}_t}(\mathbf{x}_i)$$

3. Obtenim $f_T(\mathbf{x})$

8.5 Validació

8.5.1 Validació creuada (*cross-validation*)

La tècnica de la validació creuada o *cross-validation* s'utilitza per optimitzar els paràmetres dels models predictius de forma que es redueixi el sobreajust (*overfitting*) i la predicció s'ajusti al màxim possible a les dades d'entrenament.

Les dades disponibles es divideixen aleatòriament en tres subconjunts. El conjunt d'entrenament serà per construir el model, el conjunt de validació serà per ajustar els paràmetres i el conjunt test serà per avaluar el comportament predictiu, és a dir, amb dades que no s'han introduït per construir el model (*out-of-sample*).

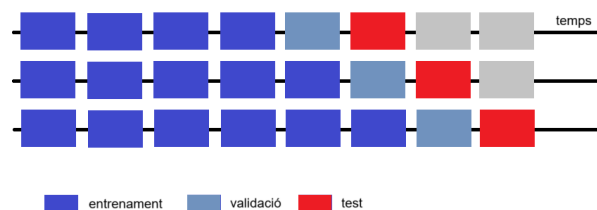
En aquest tipus de validació, se solen introduir k iteracions o *k-folds*, de manera que el conjunt de dades es divideix en k subconjunts. Per a cada iteració se selecciona un subconjunt diferent com a conjunt de prova o test i els $k - 1$ restants com a conjunt d'entrenament. Un cop fetes les k iteracions, es calcula la mitjana aritmètica dels resultats obtinguts per a cada iteració o *fold* i aquest n'és l'únic resultat final.

Els paràmetres s'escullen tenint en compte quins minimitzen l'error en la predicció de les dades del conjunt de validació. Tanmateix, aquesta tècnica assumeix que les dades són independents i distribuïdes de manera idèntica (i.i.d.), cosa que no passa amb les dades de què disposem, ja que són sèries temporals.

8.5.2 Walk-forward validation

Per això cal considerar una altra forma de validar i optimitzar el model. De l'anglès *walk-forward*, que es pot traduir al català literalment com ‘caminar endavant’, la validació *walk-forward* permet tenir en compte la dependència temporal de les dades financeres.

En aquest context, es divideixen les dades en blocs que cobreixen el mateix període de temps on les dades d'entrenament consisteixen en blocs contigus, començant des de l'inici temporal. Els blocs de validació i test es troben temporalment després dels blocs d'entrenament. El concepte de *walk-forward* es basa en el fet que per a cada *fold* o pas, hi ha blocs que no s'introdueixen a la validació, però a mesura que augmentem el pas, augmentem la longitud o durada dels blocs de tal forma que anem introduint dades als blocs d'entrenament així com als blocs de validació i test. D'aquesta forma, en l'últim *fold* totes les dades de la sèrie temporal estaran en algun bloc.



(a) *3-fold* validació *walk-forward*

Finalment, analitzem l'error en cada pas o *fold* entre el grup de validació i les prediccions i en fem la mitjana. L'objectiu doncs és trobar els paràmetres que minimitzen aquest error.

Per tal d’evitar *information leaks* (fuites d’informació), un cop s’han dividit en blocs les dades, es duu a terme un procés de depuració o *purging*, de Pardo (2018) [19]. Aquest essencialment consisteix a eliminar dades pròximes a les fronteres dels blocs o bé dades d’un mateix moment s’assignen al mateix bloc.

9 Avaluació

Un cop implementats els diferents algorismes de predicció i refnats els paràmetres que permeten el millor comportament el grup de validació, es fa servir l'error absolut mig (EAM o MAE) per mesurar el comportament predictiu dels models:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - f(\mathbf{x}_i)|, \quad (9.1)$$

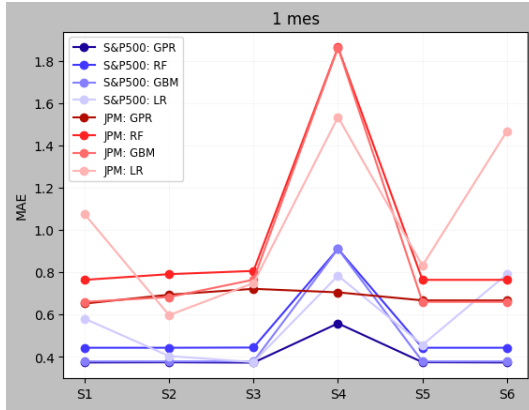
on y_i és el valor observat de IVOL i $f(\mathbf{x}_i)$ el valor predit.

A més, per demostrar la capacitat predictiva de les tècniques d'aprenentatge automàtic inclourem una regressió lineal simple (RL) de la forma

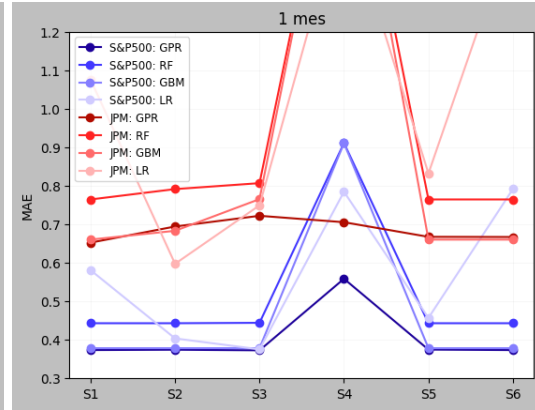
$$f_{RL}(\mathbf{x}_i) = \beta_0 + \beta \mathbf{x}_i \quad (9.2)$$

amb β de dimensió p .

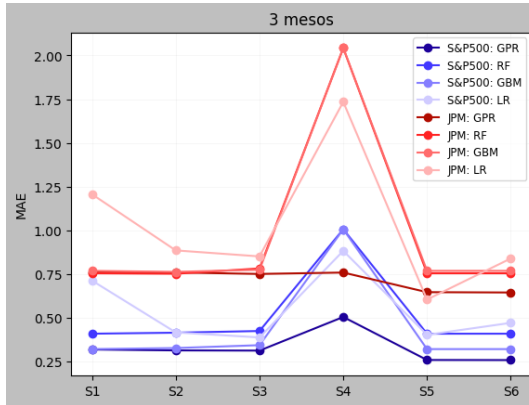
Les figures 13a - 17b mostren els resultats del MAE per a cada maduresa T avaluada per ambdós actius analitzats. **S&P 500** amb tonalitats de blau i **JPM** amb tonalitats de vermell. Per a cada actiu, es mostra el MAE obtingut amb cada algorisme emprat, Regressió del Procés Gaussià (GPR), Bosc Aleatori (RF), Potenciació del Gradient (GBM) i Regressió Lineal (LR).



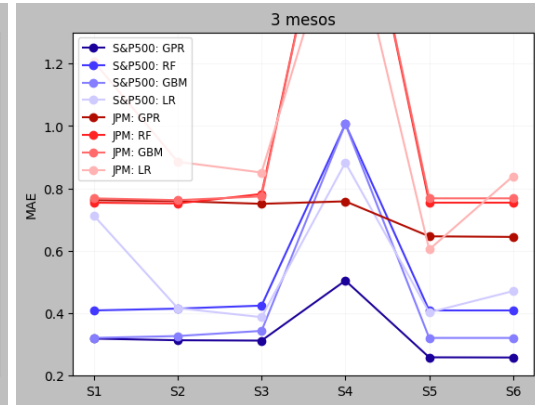
(a) MAE per a $T = 1$ mes



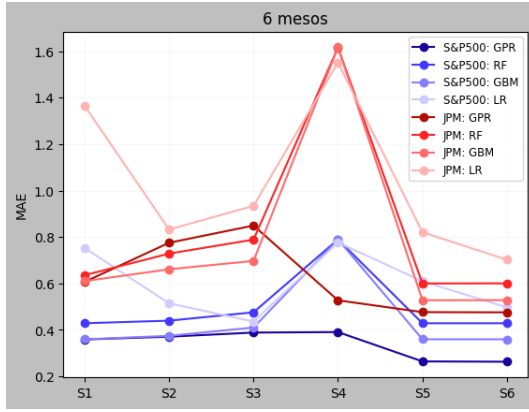
(b) MAE per a $T = 1$ mes en detall



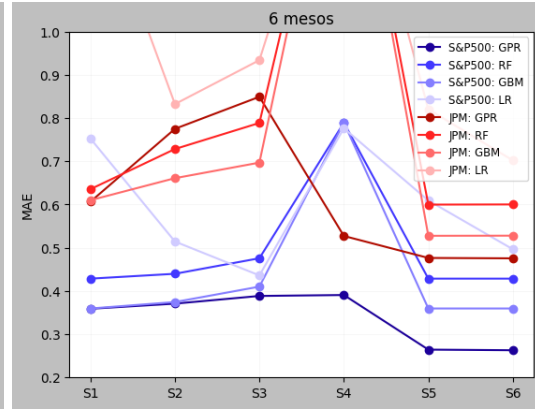
(a) MAE per a $T = 3$ mesos



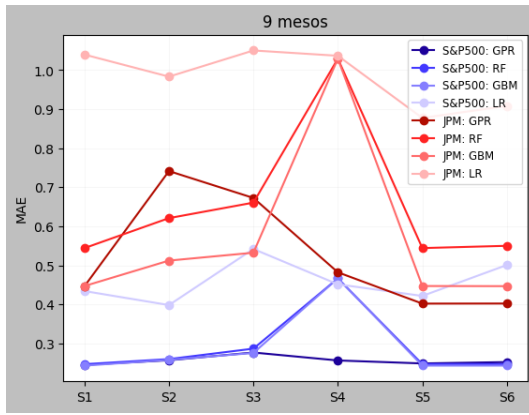
(b) MAE per a $T = 3$ mesos en detall



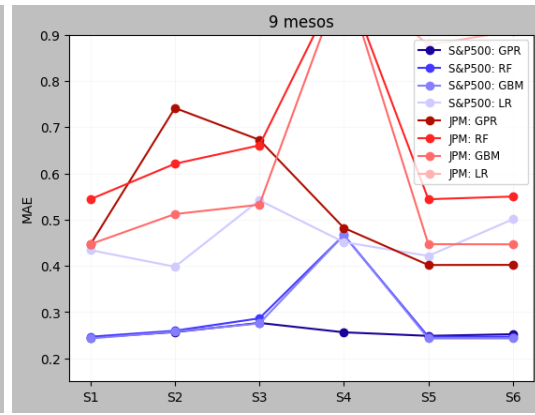
(a) MAE per a $T = 6$ mesos



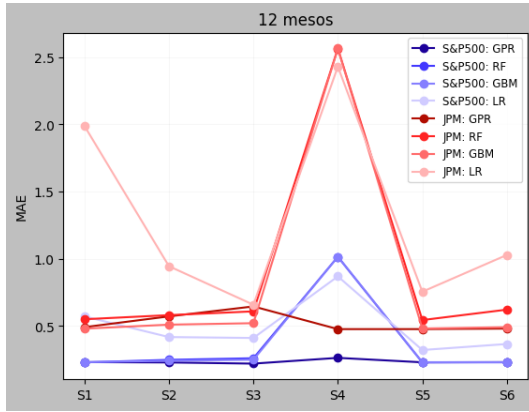
(b) MAE per a $T = 6$ mesos en detall



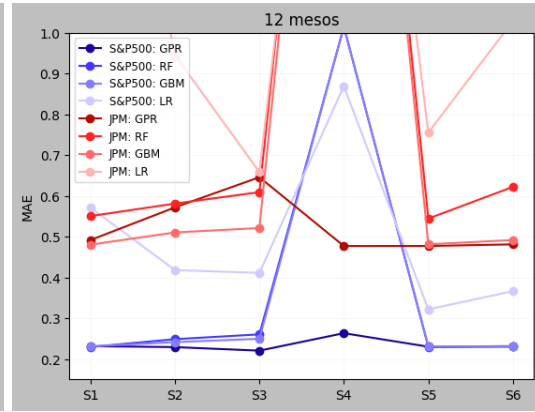
(a) MAE per a $T = 9$ mesos



(b) MAE per a $T = 9$ mesos en detall



(a) MAE per a $T = 12$ mesos



(b) MAE per a $T = 12$ mesos en detall

Figura 17: MAE de les prediccions per al S&P 500 i JPM i diferents T

Els resultats mostrats corresponen al millor conjunt de paràmetres, és a dir, que reduïxen al màxim l'error de predicció. D'aquests resultats se'n poden fer una sèrie d'observacions.

En primer lloc, i a causa de la construcció del conjunt S4, que per definició és un valor constant per a tots dos actius i durant tots els períodes, la capacitat predictiva és baixa i només la Regressió del Procés Gaussià s'adapta satisfactòriament a aquesta circumstància.

En segon lloc, observem com la introducció de l'asimetria (S5) millora dràsticament l'actuació dels algorismes RF, GBM i LR. Aquesta millora és relativa en el cas del GPR ja que, com s'acaba de comentar, la predicció per a S4 ja era prou bona, i més accentuada en el **JPM** que en el **S&P 500**. En contrast a l'anterior, la introducció de característiques com l'asimetria i la curtosi als conjunts S2 i S3 no es veu reflectit en una millora de la capacitat predictiva dels models. De fet, per algunes madureses les variables introduïdes arriben a confondre els algorismes i empitjorar la predicció comparada amb el conjunt que conté únicament una variable basada en la volatilitat, S1, sobretot en **JPM**. Això pot ser degut a l'elevada correlació que hi ha entre les variables dins de cada conjunt.

En aquesta línia, per a GPR, RF i GMB, la introducció de mesures de curtosi a S6 no millora el rendiment obtingut per S5 però, tanmateix, no en deteriora la capacitat predictiva per a cap de les madureses. Només s'observa que hi ha variacions de S6 respecte S5 per a la Regressió Lineal, però aquestes variacions empitjoren o milloren el resultat de S5 en funció de la maduresa.

En termes d'algorismes, tant GPR, RF com GBM tenen un comportament essencialment similar (menys en S4 comentat anteriorment) amb un lleuger millor comportament de la Regressió del Procés Gaussià en el cas del **S&P 500** en tots els conjunts. Així mateix, per al **JPM**, la Regressió del Procés Gaussià també obté lleugerament els millors resultats en S1, S4, S5 i S6 però per a S2 i S3 per a madureses intermèdies, el Bosc Aleatori obté millors resultats.

Per altra banda, la Regressió Lineal té resultats més irregulars. Per a ambdós actius la introducció de l'asimetria (S2 i S4) suposa millors rendiments però la curtosi (S3 i S6) no només a vegades no millora sinó que empitjora les prediccions. Aquestes fluctuacions són molt més accentuades en el cas del **JPM**. En tot cas, per a ambdós actius i totes les madureses, la Regressió Lineal és la que obté pitjors resultats a excepció d'un sol cas que no es considerarà significatiu.

Per últim, s'observa com la predicció del **JPM** obté MAEs més elevats que la del **S&P 500** per a tots els conjunts de dades i tots els algorismes de predicció.

10 Conclusions

Aquest treball s'ha dividit en dues parts clarament diferenciades i ha explorat en profunditat el paper dels *swaps* de volatilitat topada com a instruments financers essencials per a l'especulació i la gestió de riscos.

En la primera part s'ha aprofundit en el model de valoració d'opcions de Black-Scholes i en tots els resultats estocàstics rellevants del qual aquest depèn, com el moviment Brownià, pilar fonamental d'aquest model i de molts altres, dins i fora del món financer. Seguidament, i a falta de dades reals, hem hagut de recolzar-nos en nombroses ocasions de la senzillesa que presenta el model per poder obtenir tots els components necessaris per a dur a terme la modelització, tant per construir el mètode dels moments implícits del mercat com per determinar la volatilitat implícita.

Tanmateix, aquest model, encara molt present en les finances d'avui dia, presenta una sèrie de limitacions que s'han pogut observar a la secció 5 i que fa que en mercats complexos, no es capturin adequadament les dinàmiques de volatilitat.

La segona part, ha estat enfocada a la integració dels mètodes predictius en l'anàlisi i valoració de la volatilitat dels *swaps* de volatilitat topada. Els mètodes emprats, la Regressió del Procés Gaussià, el Bosc Aleatori, la Potenciació del Gradient així com la ja coneguda Regressió Lineal, ha permès introduir-nos en un camp en què només s'hi havia passat superficialment durant la carrera de matemàtiques. Gràcies a llibreries implementades de Python, es van poder fer tots els càlculs i les prediccions presentades durant el treball.

Els mètodes d'aprenentatge automàtic obren noves possibilitats al món financer i permeten que, combinats amb els models teòrics, millorin la seva aplicabilitat i precisió en els mercats cada vegada més dinàmics.

En aquest sentit, les observacions detallades en l'apartat anterior, i tot i les limitacions amb què ens hem anat trobant per accedir a dades reals del mercat, determinen un possible camí a seguir per tal de, primer, verificar que els preus dels contractes *swap* de volatilitat són coherents i segon, establir una eina per predir el comportament tant dels preus com les volatilitats.

Així doncs, s'ha vist com hi ha mètodes que s'ajusten millor que altres en termes generals (com la Regressió del Procés Gaussià) però que per a diferents actius subjacents i diferents períodes de maduresa, la decisió del mètode òptim ha de basar-se en un previ estudi sobre aquest actiu i el tipus de variables que s'hi introduiran. En termes generals hem pogut observar com per madureses elevades, els conjunts de dades S5 i S6 (S4 en la Regressió del Procés Gaussià) milloren lleugerament els resultats de S1, S2, i S3 per ambdós actius.

Tanmateix, donades les possibles limitacions del model de Black-Scholes en les construccions dels preus de les opcions (com que la volatilitat sigui en tots els períodes constant per hipòtesi) fa que, en addició a l'estudi realitzat, els resultats amb preus reals puguin ser lleugerament diferent i, per tant, reforçar les conclusions extretes en el present treball com descartar-les. Recolzant-nos en un article de Höcht, Schoutens i Verschuere (2024) [16] concloem que aquestes, lluny de descartar-se, es reforcen i, en conseqüència, observem que la construcció teòrica el que fa és homogeneïtzar i igualar els resultats, llimar-ne les diferències i, en conseqüència, obtenir unes previsions més parelles per a tots els conjunts de dades, però en cap cas contradir els resultats que s'obtenen amb dades reals de preus.

Per últim, aquest treball ha mostrat com les eines d'aprenentatge automàtic poden complementar l'enfocament tradicional, oferint noves perspectives per a analistes. Aquesta dualitat entre enfocaments tradicionals i tècniques modernes no ha de ser vista com un reemplaçament, sinó com una sinergia, ja que, només integrant ambdós punts de vista serà possible afrontar els desafiaments en un entorn en constant evolució. En l'equilibri de les dues hi resideix el futur dels mercats financers, tal com hem pogut aprendre de l'incident del **Flash Crash**.

Referències

- [1] Bachelier, L. (1900). *Théorie de la spéculation*. Annales Scientifiques de l'Ecole Normale Supérieure. Quatrième Serie, 17, 21-86. <https://doi.org/10.24033/asens.476>
- [2] Bakshi, G., Madan, D. (2000). *Spanning and derivative-security valuation*. Journal of Financial Economics, 55(2), 205-238. [https://doi.org/10.1016/s0304-405x\(99\)00050-1](https://doi.org/10.1016/s0304-405x(99)00050-1)
- [3] Black, F., Scholes, M. (1973). *The pricing of options and corporate liabilities*. The Journal of Political Economy, 81(3), 637-654. <https://doi.org/10.1086/260062>
- [4] Breiman, L. (1996). *Bagging predictors*. Machine Learning, 24(2), 123-140. <https://doi.org/10.1007/bf00058655>
- [5] Breiman, L. (2001). *Random Forests*. Machine Learning, 45(1), 5-32.
- [6] Breiman, L., Friedman, J., Stone, C. J., Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman & Hall/CRC.
- [7] Brenner, M., Subrahmanyam, M. G. (1988). *A simple formula to compute the implied standard deviation*. Financial Analysts Journal, 44(5), 80-83. <https://doi.org/10.2469/faj.v44.n5.80>
- [8] Capinski, M., Kopp, E., Traple, J. (2012). *Stochastic calculus for finance*. Cambridge University Press.
- [9] Carr, P., Madan, D. (2001). *Towards a theory of volatility trading*. En E. Jouini, J. Cvitanic, M. Musiela (Eds.), Option Pricing, Interest Rates and Risk Management (pp. 458-476). Cambridge University Press.
- [10] Corrado, C. J., Miller, T. W., Jr. (1996). *A note on a simple, accurate formula to compute implied standard deviations*. Journal of Banking and Finance, 20(3), 595-603. [https://doi.org/10.1016/0378-4266\(95\)00014-3](https://doi.org/10.1016/0378-4266(95)00014-3)
- [11] Demeterfi, K., Derman, E., Kamal, M., Zou, J. (1999). *More Than You Ever Wanted to Know About Volatility Swaps*. Goldman Sachs Quantitative Strategies Research Notes.
- [12] Flavell, R. R. (2012). *Swaps and Other Derivatives (2a ed.)*. John Wiley & Sons.
- [13] Friedman, J. H. (2002). *Stochastic gradient boosting*. Computational Statistics & Data Analysis, 38(4), 367-378. [https://doi.org/10.1016/s0167-9473\(01\)00065-2](https://doi.org/10.1016/s0167-9473(01)00065-2)
- [14] Glasserman, P. (2014). *Monte Carlo Methods in Financial Engineering*. Springer.
- [15] Harrison, J. M., Pliska, S. R. (1981). *Martingales and stochastic integrals in the theory of continuous trading*. Stochastic Processes and Their Applications, 11(3), 215-260. [https://doi.org/10.1016/0304-4149\(81\)90026-0](https://doi.org/10.1016/0304-4149(81)90026-0)
- [16] Höcht, S., Schoutens, W., Verschuere, E. (2024). *On the pricing of capped volatility swaps using machine learning techniques*. Quantitative Finance, 1-14. <https://doi.org/10.1080/14697688.2024.2305643>
- [17] Karatzas, J., Shreve, S. (1988). *Brownian motion and stochastic calculus*. Springer.

- [18] Lamberton, D., Lapeyre, B. (2007). *Introduction to Stochastic Calculus Applied to Finance (2nd ed.)*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781420009941>
- [19] Lopez de Prado, M. (2018). The 10 reasons most machine learning funds fail. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3104816>
- [20] Madan, D., Schoutens, W. (2016). *Applied Conic Finance*. Cambridge University Press.
- [21] Mixon, S. (2011). *What does implied volatility skew measure?* The Journal of Derivatives, 18(4), 9-25. <https://doi.org/10.3905/jod.2011.18.4.009>
- [22] Rasmussen, C. E., Williams, C. K. (2006). *Gaussian processes for machine learning*. Mit Press.
- [23] Ross, S. A. (1976). *The arbitrage theory of capital asset pricing*. Journal of Economic Theory, 13(3), 341-360. [https://doi.org/10.1016/0022-0531\(76\)90046-6](https://doi.org/10.1016/0022-0531(76)90046-6)
- [24] Sanz-Sole, M. (1999) *Probabilitats*.