

Empresas de petróleo

Introdução

A precificação e otimização de carteira de ações é um tema de grande importância no campo da finanças, e tem recebido atenção considerável, especialmente quando se trata de empresas do setor de petróleo listadas na bolsa. A volatilidade dos preços do petróleo, a influência de fatores macroeconômicos e a dinâmica do mercado de energia tornam essas empresas alvos de análises detalhadas e estratégias de investimento específicas.

Este trabalho tem como objetivo principal desenvolver um modelo para a precificação de empresas do ramo de petróleo listadas na bolsa. Para atingir esse objetivo, adotaremos uma abordagem que considera o preço do barril de petróleo como um dos principais determinantes dos preços das ações. Além disso, será incorporado um componente autoregressivo nos preços e também o índice.

Em seguida, será realizada estimativa da volatilidade das ações das empresas de petróleo. Para isso, utilizaremos modelos GARCH (Generalized Autoregressive Conditional Heteroskedasticity) para cada ação e cada variável relevante. Isso nos permitirá criar uma matriz de covariância que captura as relações de volatilidade entre as empresas e as variáveis que as influenciam.

A matriz de covariância é uma ferramenta essencial no cálculo da melhor carteira de investimentos. Com base nessa matriz e nas expectativas de retorno, o próximo passo será montar uma carteira que maximize o índice de Sharpe. O índice de Sharpe é uma medida fundamental que considera o retorno esperado e a volatilidade da carteira.

Apresentação dos dados

Neste trabalho, os dados utilizados compreendem o histórico de preços do barril de petróleo Brent, os dados de preços do índice IBOV (Índice Bovespa), bem como o histórico de preços de todas as ações pertencentes ao setor de petróleo, gás e biocombustíveis.

A seguir, apresenta-se uma tabela descritiva dos dados, que são os preços de fechamento ajustados desde o ano de 2010 até o dia 31 de outubro.

Tabela 1: Descrição dos dados.

	Obs.	Média	Desv.Pad	Min	25%	50%	75%	Máx
Brent	4031.0	78.56	25.17	19.33	58.24	75.91	102.74	146.08
RRRP3	713.0	36.59	5.63	20.80	32.35	36.03	40.84	51.31
CSAN3	3872.0	8.47	5.69	1.23	4.02	6.63	11.59	24.20
ENAT3	3041.0	7.89	4.42	1.86	4.16	6.90	10.77	23.24
PETR3	3872.0	10.85	6.30	2.28	6.48	9.50	12.76	41.56
PETR4	3872.0	9.12	5.52	1.49	5.87	7.81	10.44	38.52
PRI03	3110.0	10.45	12.86	0.14	0.93	3.69	18.89	50.20
RPMG3	3872.0	8.34	6.53	1.52	3.36	5.60	11.98	32.34
VBBR3	1358.0	18.32	3.24	11.18	15.96	18.06	20.58	26.85
UGPA3	3872.0	17.72	8.88	5.01	11.38	18.15	23.60	36.45
LUPA3	3872.0	1369.74	2523.17	0.89	2.60	6.56	1520.85	10761.76
OPCT3	656.0	4.25	2.27	1.62	2.69	3.18	5.19	11.15
IBOV	3863.0	74054.49	24458.07	29435.00	54881.00	65051.00	97871.00	130776.00

A seguir, está a matriz de correlação das variáveis durante todo o período. Para facilitar a visualização, ela está dividida em duas tabelas distintas.

Tabela 2.1: Matriz de correlação (Parte I).

	Brent	RRRP3	CSAN3	ENAT3	PETR3	PETR4	PRI03
Brent	1.00	0.28	-0.12	0.26	0.25	0.24	0.36
RRRP3	0.28	1.00	0.39	0.45	-0.23	-0.24	-0.07
CSAN3	-0.12	0.39	1.00	0.82	0.48	0.57	0.51
ENAT3	0.26	0.45	0.82	1.00	0.80	0.80	0.79
PETR3	0.25	-0.23	0.48	0.80	1.00	0.99	0.83
PETR4	0.24	-0.24	0.57	0.80	0.99	1.00	0.81
PRI03	0.36	-0.07	0.51	0.79	0.83	0.81	1.00
RPMG3	0.38	0.28	-0.59	-0.28	-0.08	-0.18	-0.05
VBBR3	-0.02	0.32	0.52	0.29	-0.09	-0.06	-0.01
UGPA3	-0.37	-0.14	0.13	-0.69	-0.44	-0.35	-0.70
LUPA3	0.28	0.42	-0.48	-0.01	0.18	0.05	0.36
OPCT3	-0.62	0.27	0.43	-0.18	-0.34	-0.28	-0.24
IBOV	-0.08	0.41	0.91	0.85	0.62	0.68	0.57

Tabela 2.2: Matriz de correlação (Parte II).

	RPMG3	VBBR3	UGPA3	LUPA3	OPCT3	OSXB3	IBOV
Brent	0.38	-0.02	-0.37	0.28	-0.62	0.46	-0.08
RRRP3	0.28	0.32	-0.14	0.42	0.27	0.06	0.41
CSAN3	-0.59	0.52	0.13	-0.48	0.43	-0.46	0.91
ENAT3	-0.28	0.29	-0.69	-0.01	-0.18	-0.05	0.85
PETR3	-0.08	-0.09	-0.44	0.18	-0.34	-0.07	0.62
PETR4	-0.18	-0.06	-0.35	0.05	-0.28	-0.12	0.68
PRI03	-0.05	-0.01	-0.70	0.36	-0.24	0.29	0.57
RPMG3	1.00	0.28	-0.47	0.72	0.62	0.73	-0.44
VBBR3	0.28	1.00	0.04	0.38	0.42	0.12	0.65
UGPA3	-0.47	0.04	1.00	-0.68	0.80	-0.56	-0.04
LUPA3	0.72	0.38	-0.68	1.00	0.17	0.70	-0.31
OPCT3	0.62	0.42	0.80	0.17	1.00	0.79	0.65
IBOV	-0.44	0.65	-0.04	-0.31	0.65	-0.29	1.00

Os dados foram obtidos por meio da API do Yahoo Finance para Python (YFinance) no dia 10 de novembro de 2023.

Modelos autoregressivos

Nesta fase do projeto, será considerada a aplicação de modelos autoregressivos. Com esse propósito, é apresentada a seguir as tabelas contendo os valores da Função de Autocorrelação (ACF) e da Função de Autocorrelação Parcial (PACF) das variáveis previamente exploradas.

Tabela 3: PACF para a primeira diferença.

	0	1	2	3	4	5	6	7	8	9	10
Brent	1.0	-0.02	0.00	0.00	0.03	-0.02	-0.03	-0.02	-0.02	0.04	0.02
RRRP3	1.0	0.14	-0.05	0.02	0.01	-0.06	0.02	0.04	0.00	0.03	-0.01
CSAN3	1.0	-0.01	-0.00	-0.03	-0.03	0.02	-0.02	-0.01	-0.04	-0.02	0.01
ENAT3	1.0	-0.06	-0.01	-0.03	0.03	-0.02	0.00	0.02	0.01	-0.00	0.00
PETR3	1.0	-0.02	-0.01	0.00	0.03	0.02	-0.01	0.04	-0.03	0.01	0.02
PETR4	1.0	-0.02	-0.00	0.00	0.03	0.02	-0.01	0.02	-0.01	0.01	0.03
PRI03	1.0	0.04	-0.01	-0.03	-0.03	0.00	-0.01	0.03	0.06	0.00	0.01
RPMG3	1.0	-0.05	0.04	0.03	-0.05	0.00	-0.00	-0.00	0.02	-0.00	-0.01
VBBR3	1.0	-0.02	0.02	-0.01	-0.03	0.03	-0.12	0.04	-0.01	0.05	0.00
UGPA3	1.0	-0.11	0.02	-0.01	-0.02	0.04	-0.09	0.06	-0.01	0.03	0.01
LUPA3	1.0	-0.05	-0.06	0.04	0.04	-0.02	-0.05	-0.03	0.03	0.05	0.01
OPCT3	1.0	0.11	-0.07	-0.01	0.03	-0.02	-0.07	-0.01	-0.02	0.08	0.06
IBOV	1.0	-0.06	-0.00	-0.01	-0.00	-0.01	-0.06	0.03	-0.03	0.04	0.02

Tabela 4: ACF para a primeira diferença.

	0	1	2	3	4	5	6	7	8	9	10
Brent	1.0	-0.02	0.00	0.00	0.03	-0.02	-0.02	-0.02	-0.02	0.04	0.02
RRRP3	1.0	0.14	-0.02	0.01	0.02	-0.06	0.00	0.04	0.01	0.02	0.00
CSAN3	1.0	-0.01	-0.00	-0.03	-0.03	0.02	-0.02	-0.00	-0.04	-0.02	0.01
ENAT3	1.0	-0.06	-0.00	-0.03	0.03	-0.02	0.00	0.02	0.01	-0.00	0.00
PETR3	1.0	-0.02	-0.01	0.00	0.03	0.02	-0.02	0.04	-0.03	0.01	0.01
PETR4	1.0	-0.02	0.00	0.00	0.03	0.02	-0.01	0.02	-0.01	0.02	0.03
PRI03	1.0	0.04	-0.01	-0.03	-0.03	0.00	-0.01	0.03	0.07	0.01	0.01
RPMG3	1.0	-0.05	0.04	0.03	-0.05	0.01	-0.01	-0.01	0.03	-0.01	-0.01
VBBR3	1.0	-0.02	0.02	-0.01	-0.03	0.03	-0.13	0.05	-0.02	0.05	0.01
UGPA3	1.0	-0.11	0.03	-0.02	-0.02	0.04	-0.09	0.08	-0.03	0.04	0.00
LUPA3	1.0	-0.05	-0.05	0.05	0.03	-0.03	-0.05	-0.02	0.04	0.04	-0.00
OPCT3	1.0	0.11	-0.06	-0.02	0.03	-0.01	-0.08	-0.02	-0.01	0.07	0.08
IBOV	1.0	-0.06	0.00	-0.01	-0.00	-0.01	-0.06	0.04	-0.04	0.04	0.02

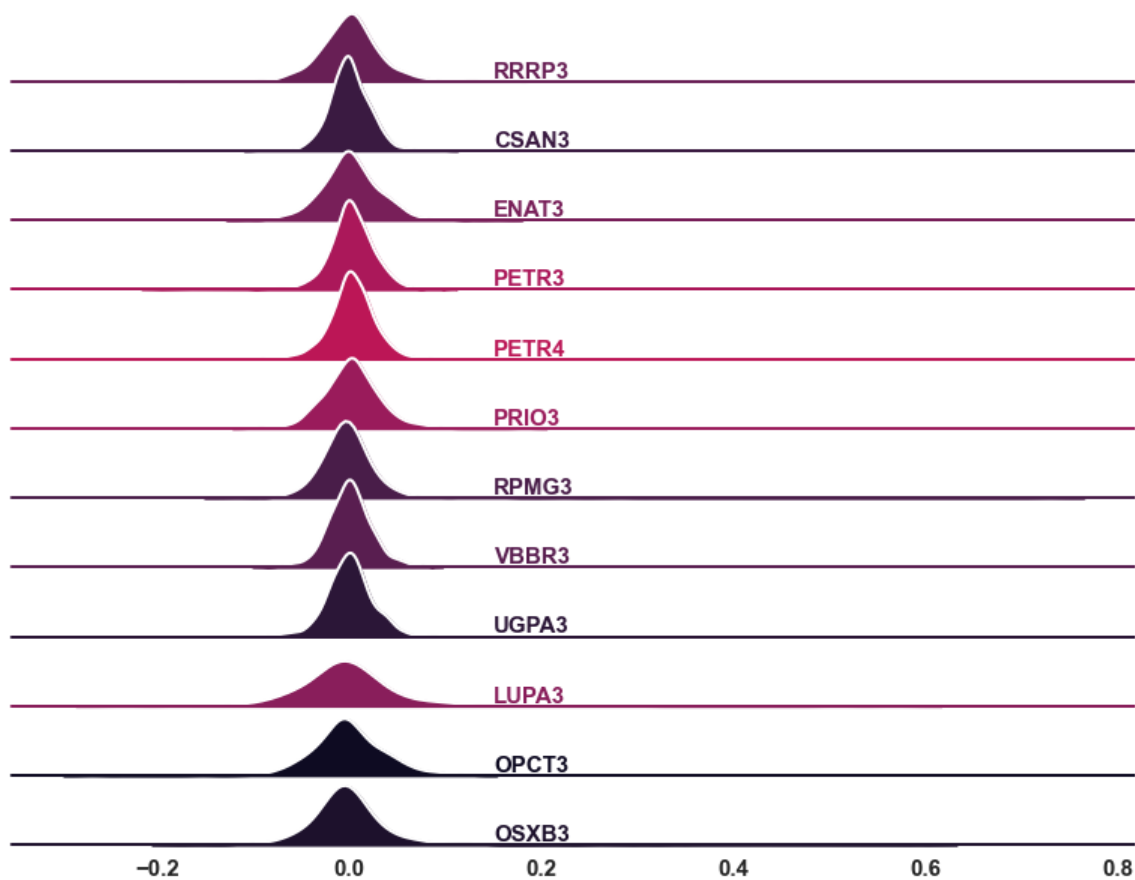
Apesar de não ter obtido resultados muito satisfatórios ao utilizar o modelo autoregressivo, ainda é uma abordagem valiosa, uma vez que possibilita encontrar o desvio padrão do erro. Isso, por sua vez, abre caminho para a construção posterior da matriz de covariância teórica.

De forma generalizada, cada ação será modelada como uma função linear que incorpora componentes autoregressivos e regressivos, envolvendo as variáveis de 1 até 3 lags. Assim, para cada variação de preço, consideramos até 7 variáveis explicativas. A seleção do modelo mais apropriado será baseada no critério de informação AIC (Critério de Informação de Akaike). Este processo envolverá a avaliação de todas as combinações possíveis de variáveis explicativas, e o modelo que minimizar o valor do AIC será escolhido.

Na tabela a seguir, apresentam-se os resultados das simulações. O retorno diário de cada ação é modelado como a soma de processos AR(3) de outras empresas, incluindo o índice e o preço do barril de petróleo. É importante destacar que o modelo foi treinado no período de janeiro de 2010 a dezembro de 2022, e as estatísticas a seguir foram calculadas usando como base o ano de 2023.

Tabela 5: Modelos selecionados.

Ativo	Regressores	AIC	BIC	R ²	Acerto de sinal
RRRP3	7	-2741.03	-2704.90	0.05	0.54
CSAN3	6	-3282.13	-3250.52	0.03	0.52
ENAT3	5	-2828.70	-2801.60	0.03	0.56
PETR3	4	-3073.02	-3050.44	0.02	0.54
PETR4	5	-3079.36	-3052.27	0.02	0.52
PRI03	3	-2862.67	-2844.61	0.02	0.51
RPMG3	6	-2352.06	-2320.45	0.04	0.52
VBBR3	1	-3193.91	-3184.87	0.00	0.51
UGPA3	7	-3093.37	-3057.24	0.03	0.54
LUPA3	5	-1847.90	-1820.80	0.04	0.56
OPCT3	7	-2515.10	-2478.97	0.04	0.55

Gráfico 1: Distribuição dos Resíduos.

Modelos de volatilidade

O modelo anterior permitia a construção de expectativas de preços para o dia seguinte. No entanto, o modelo a seguir tem como objetivo modelar a covariância futura. Realizando manipulações algébricas simples na definição tautológica de covariância, podemos escrever da seguinte forma:

$$\text{COV}(X, Y) = \frac{V(X + Y) - V(X) - V(Y)}{2} \quad (1)$$

$$\text{COV}(X, Y) = \frac{-V(X - Y) + V(X) + V(Y)}{2} \quad (2)$$

$$\text{COV}(X, Y) = \frac{V(X + Y) - V(X - Y)}{4} \quad (3)$$

É evidente que a covariância é, na realidade, uma função de 2 variâncias distintas. Portanto, o objetivo neste contexto é desenvolver um modelo para cada uma dessas variâncias individuais e, subsequentemente, combinar essas equações para calcular a covariância. A seguir, apresentamos uma visualização das autocorrelações e autocorrelações parciais dos resíduos quadrados do modelo AR do tópico anterior:

Tabela 6: PACF dos resíduos quadrados.

	0	1	2	3	4	5	6	7	8	9	10
RRRP3	1.0	0.20	-0.07	0.03	-0.01	0.11	0.01	-0.03	0.05	0.02	-0.02
CSAN3	1.0	0.18	0.09	0.15	0.11	0.07	0.01	0.11	0.06	0.04	0.03
ENAT3	1.0	0.03	-0.00	0.05	-0.02	-0.00	0.01	0.01	0.02	-0.01	0.00
PETR3	1.0	0.21	0.05	0.11	0.01	0.23	0.19	-0.07	0.07	0.07	0.07
PETR4	1.0	0.23	0.03	0.11	0.02	0.16	0.16	-0.06	0.06	0.03	0.05
PRI03	1.0	0.08	-0.05	-0.02	-0.06	-0.03	0.02	0.01	-0.02	0.03	0.01
RPMG3	1.0	0.46	-0.23	0.12	-0.07	0.02	-0.03	0.01	-0.02	0.01	-0.01
VBBR3	1.0	0.03	0.02	0.00	0.06	0.02	0.08	-0.01	-0.00	0.01	0.01
UGPA3	1.0	-0.01	0.12	-0.00	0.03	0.02	-0.05	-0.00	-0.03	0.00	0.01
LUPA3	1.0	0.13	0.02	0.04	0.01	0.01	0.01	0.01	0.02	0.04	0.01
OPCT3	1.0	0.09	0.08	-0.00	0.01	-0.02	0.01	0.10	0.08	0.02	-0.06

Tabela 7: ACF dos resíduos quadrados.

	0	1	2	3	4	5	6	7	8	9	10
RRRP3	1.0	0.20	-0.02	0.01	-0.00	0.10	0.05	-0.03	0.03	0.04	-0.01
CSAN3	1.0	0.18	0.12	0.18	0.16	0.14	0.08	0.17	0.14	0.11	0.11
ENAT3	1.0	0.03	-0.00	0.05	-0.02	-0.00	0.01	0.01	0.02	-0.01	0.00
PETR3	1.0	0.21	0.09	0.13	0.06	0.24	0.27	0.04	0.12	0.14	0.14
PETR4	1.0	0.23	0.08	0.13	0.07	0.18	0.24	0.04	0.09	0.10	0.10
PRI03	1.0	0.08	-0.04	-0.03	-0.06	-0.03	0.02	0.02	-0.01	0.03	0.01
RPMG3	1.0	0.46	0.03	-0.00	-0.00	-0.01	-0.02	-0.02	-0.02	-0.01	-0.01
VBBR3	1.0	0.03	0.02	0.00	0.06	0.03	0.08	-0.00	0.00	0.01	0.02
UGPA3	1.0	-0.01	0.12	-0.01	0.04	0.02	-0.04	0.00	-0.04	0.01	-0.00
LUPA3	1.0	0.03	0.02	0.04	0.01	0.01	0.01	0.02	0.02	0.04	0.01
OPCT3	1.0	0.03	0.02	-0.00	0.01	-0.02	0.00	0.10	0.09	0.02	-0.05

Analisando as autocorrelações e autocorrelações parciais, será escolhido para generalizar os modelos GARCH com até 3 lags cada. Portanto, podemos expressar a equação de cada ϵ^2 da seguinte maneira:

$$\epsilon_t^2 = \alpha_1 \cdot \epsilon_{t-1}^2 + \alpha_2 \cdot \epsilon_{t-2}^2 + \alpha_3 \cdot \epsilon_{t-3}^2 + \beta_1 \cdot \sigma_{t-1}^2 + \beta_2 \cdot \sigma_{t-2}^2 + \beta_3 \cdot \sigma_{t-3}^2 \quad (4)$$

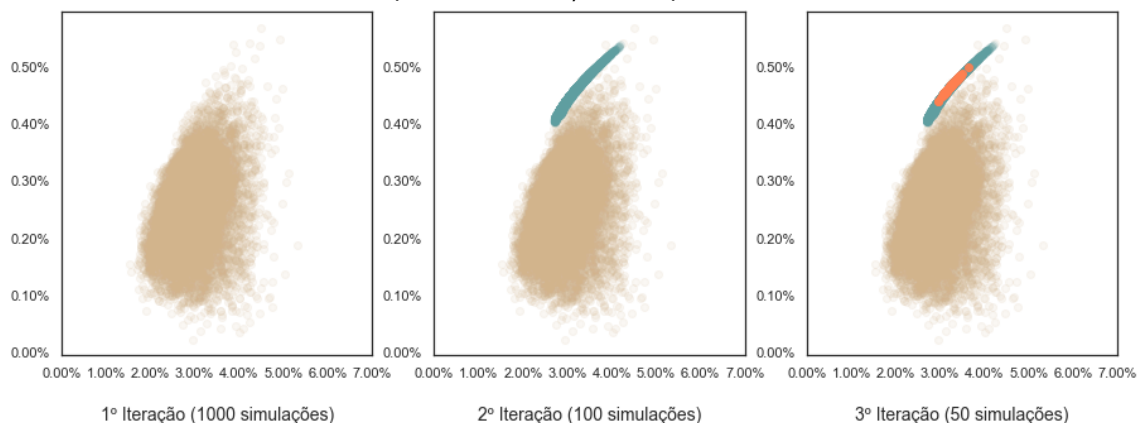
Onde α_1 , α_2 , α_3 , β_1 , β_2 , e β_3 são os coeficientes GARCH e σ_{t-1}^2 , σ_{t-2}^2 , e σ_{t-3}^2 são as variâncias condicionais nos tempos $t - 1$, $t - 2$, e $t - 3$, respectivamente. Essa estratégia será aplicada tanto para explicar a volatilidade de cada ativo como também para explicar também a volatilidade da soma de ativos.

Utilizando os modelos

Após ter criado expectativas de retorno diário das ações do ramo de petróleo e ter também criado expectativas para a matriz de covariância para cada dia. É possível, com esses inputs, aplicar o modelo de CAPM e buscar a carteira que maximiza o Sharp para cada dia.

Com o objetivo de otimizar o desempenho computacional em face da necessidade de simular mais de 200 carteiras distintas, recorremos a um conceito desenvolvido por Markowitz. Segundo essa abordagem, a carteira ótima pode ser representada como uma ponderação de duas carteiras distintas que residem na fronteira eficiente. Para simplificar as simulações, dividimos o processo em três grupos, em que cada simulação reflete uma ponderação das carteiras candidatas a fazer parte da fronteira eficiente. Esse enfoque assegura que o computador execute simulações de maneira eficiente, evitando redundâncias. O gráfico a seguir ilustra esse procedimento.

Gráfico 2: Método para a otimização computacional com Markowitz.



O gráfico a seguir nos permite comparar o desempenho da estratégia com o desempenho dos ativos que compõem a carteira. Vale destacar que os dados utilizados já consideram o reinvestimento dos dividendos. A linha **"Portfólio"** na cor vinho representa as cotas históricas da carteira, rebalanceadas diariamente (incluindo posições vendidas), levando em consideração as expectativas geradas nas etapas anteriores e com base nos parâmetros calculados entre 2010 e 2022. Por outro lado, a linha **"Best Portfólio"** em verde retrata o retorno de uma carteira estritamente comprada, rebalanceada apenas uma vez em 1º de janeiro de 2023 e utilizando informações do futuro, ou seja, uma carteira de investimentos ex-post.

Gráfico 3: Comparação da carteira com outras.

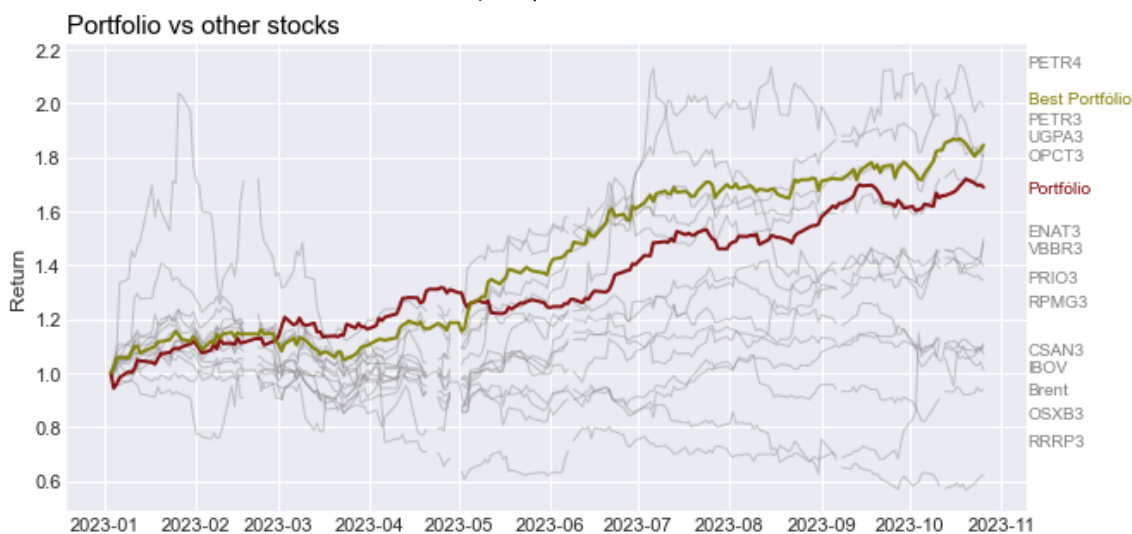


Tabela 8: Estatísticas de mercado.

	Retorno	Volatilidade	Sharpe	VaR 5%	>IBOV
RRRP3	-6.11%	3.30%	-34.60%	-4.59%	45.15%
CSAN3	10.90%	1.93%	-3.61%	-3.10%	48.54%
ENAT3	49.72%	2.67%	88.86%	-4.03%	51.46%
PETR3	82.73%	2.08%	213.89%	-3.03%	55.34%
PETR4	98.60%	2.05%	265.99%	-3.13%	57.28%
PRI03	39.94%	2.32%	75.88%	-3.48%	47.09%
RPMG3	34.36%	3.05%	46.19%	-4.18%	43.20%
VBBR3	48.28%	2.25%	101.69%	-2.96%	52.43%
UGPA3	82.34%	1.90%	232.62%	-2.44%	53.88%
LUPA3	-37.74%	4.59%	-68.30%	-6.38%	40.29%
OPCT3	80.52%	3.00%	143.98%	-4.56%	51.46%
Brent	7.10%	1.94%	-15.92%	-3.72%	53.40%
IBOV	10.19%	1.07%	-10.68%	-1.74%	0.00%
Portfólio	68.89%	1.31%	273.65%	-1.65%	57.28%
Best Portfólio	84.52%	1.43%	319.13%	-2.14%	59.71%

É notável que o portfólio gerado neste exercício apresentou resultados interessantes. Um aspecto a ser observado é que o Value at Risk (VaR) de 5% deste portfólio é menor do que o de todos os outros ativos. Isso pode ser atribuído ao fato de que este não é um portfólio estritamente comprado. Como a correlação dos eventos de cauda, ou seja, eventos 'excepcionais', tende a ser estritamente positiva em ativos da mesma categoria e do mesmo mercado, um portfólio que possui posições vendidas pode não necessariamente reduzir a expectativa de volatilidade, mas ainda assim reduzirá a expectativa de VaR. Isso ocorre porque se espera que eventos de cauda tenham correlação positiva.

Resultados

Nesse bloco, realizaremos uma análise exploratória dos resultados, incluindo comparações entre as projeções da carteira e os resultados efetivamente alcançados, juntamente com testes de hipóteses. Os gráficos a seguir apresentam o intervalo de confiança de 2 desvios padrão, com base na própria expectativa de volatilidade como parâmetro. A linha pontilhada representa a cota considerando o retorno esperado para o dia, enquanto a linha contínua retrata o retorno real.

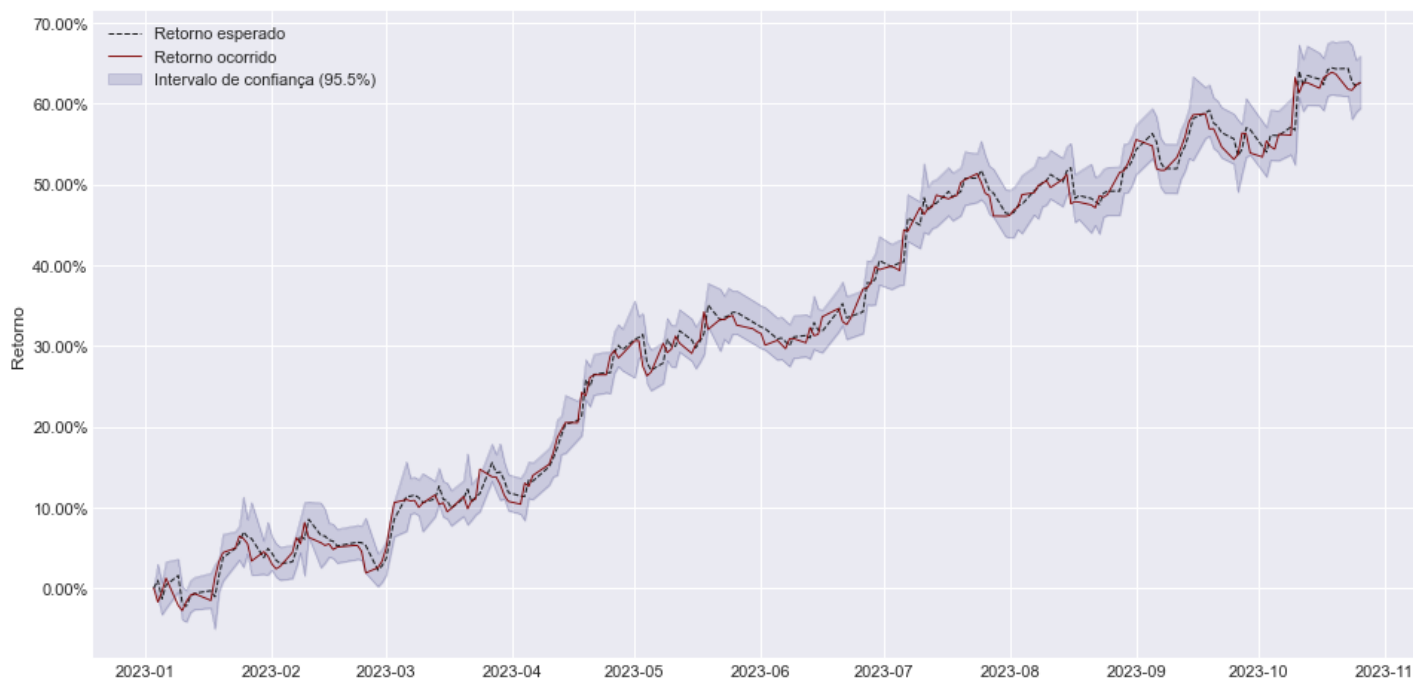
Gráfico 4: Cota com intervalo de confiança.

Gráfico 5: Intervalo de confiança sem cota.

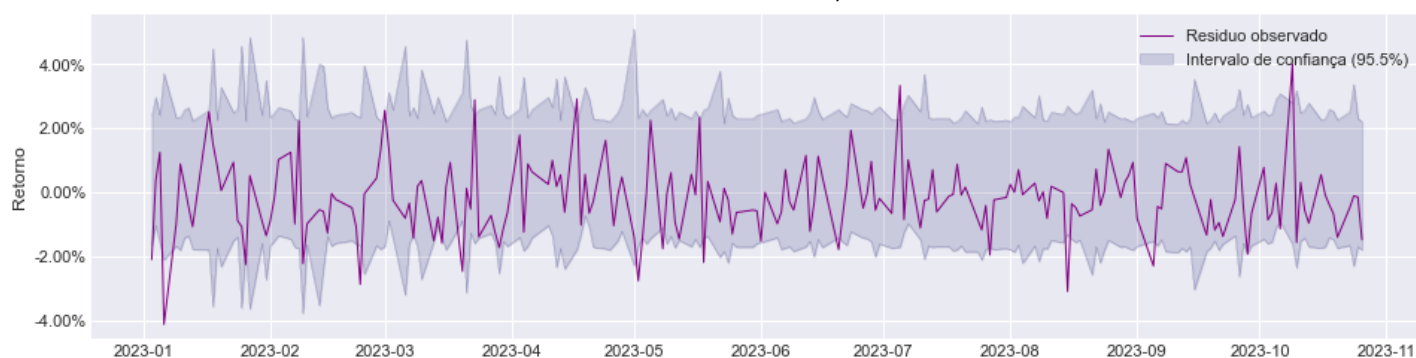
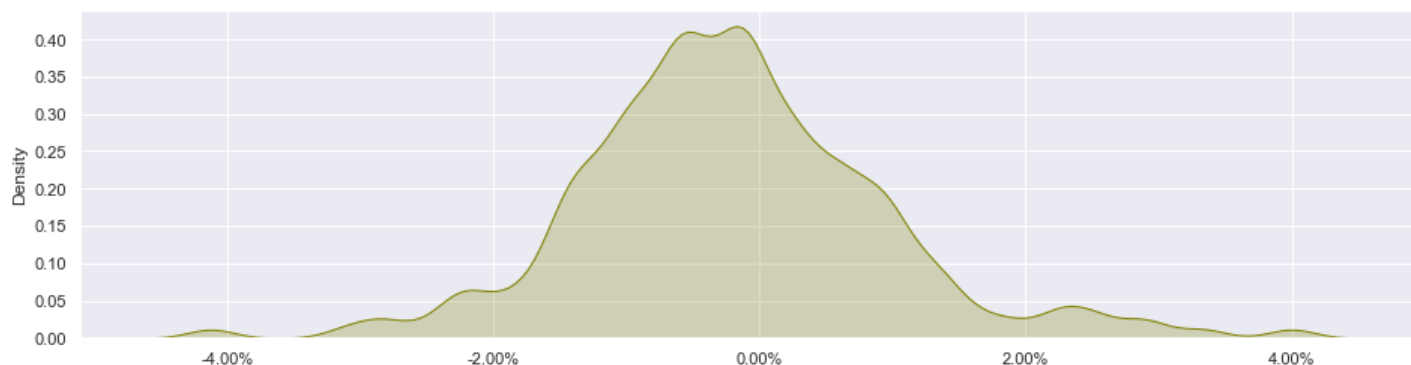
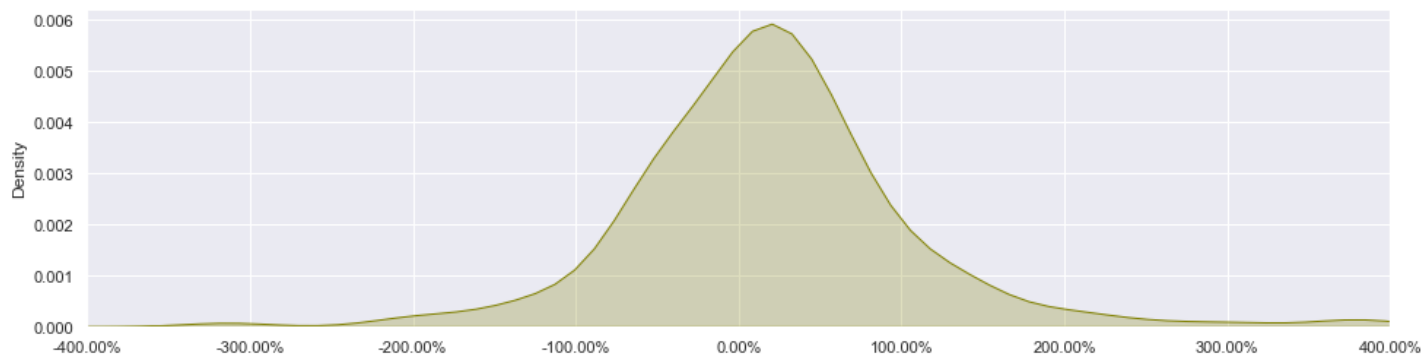


Gráfico 6: Histograma dos resíduos.



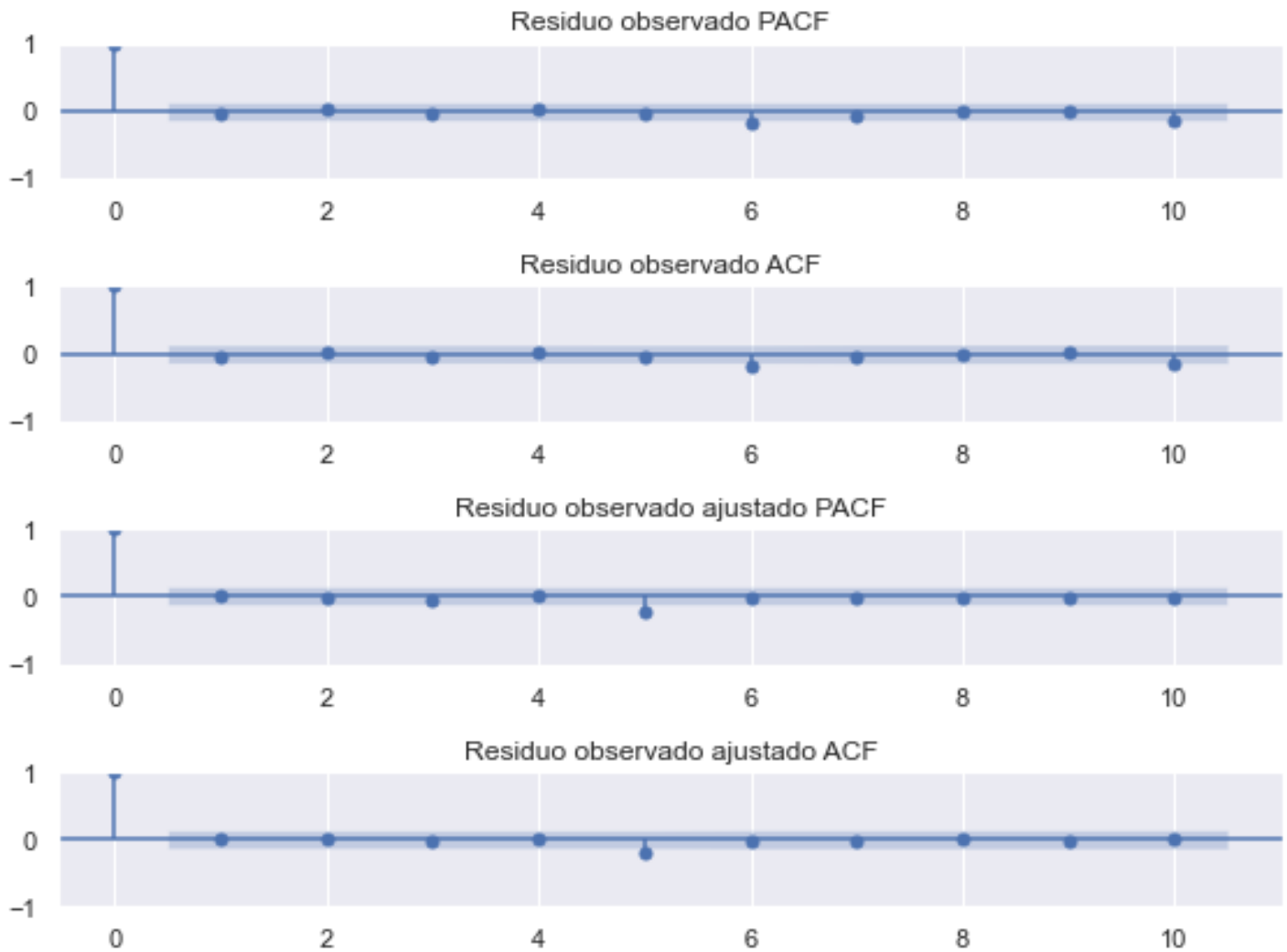
A própria existência de modelos de volatilidade sugere a presença de não-homoscedasticidade nos resíduos. Para uma melhor visualização dos resíduos, é apresentado abaixo o histograma das estatísticas t observadas. Essas estatísticas t são calculadas dividindo os resíduos pela expectativa de volatilidade. É importante ressaltar que as expectativas de volatilidade mencionadas aqui se referem, na verdade, às expectativas de volatilidade dos resíduos, não do portfólio em si.

Gráfico 7: Histograma dos resíduos ajustados.



O resíduo imediatamente após a modelagem apresenta uma aparência consideravelmente mais próxima da hipótese de normalidade dos resíduos

Gráfico 8: Autocorrelações dos resíduos do portfólio.



Como é observado acima, o resíduo do retComo evidenciado acima, os resíduos do retorno esperado exibem ausência de autocorrelação, seja de forma parcial ou global, e também apresentam uma distribuição normal quando ajustados pelas expectativas de volatilidade.

Conclusão

Este estudo econométrico demonstra que mesmo em casos com baixos coeficientes de determinação (R^2), a abordagem econométrica pode ser valiosa, uma vez que não apenas permite criar expectativas, mas também analisar o comportamento dos resíduos. O cerne deste trabalho envolveu a construção de modelos de volatilidade satisfatórios, que, por sua vez, facilitaram a criação de modelos de covariância mais robustos. Com a ajuda da Teoria Moderna de Portfólio, as expectativas de covariância e variância se mostraram ferramentas úteis para otimizar o Índice de Sharpe, resultando em retornos esperados com intervalos de confiança mais favoráveis.

t

Espero que goste!

Referências

- 1 Payal Soni, Yogya Tewari, and Prof. Deepa Krishnan (2021). *Machine Learning Approaches in Stock Price Prediction: A Systematic Review*. *Journal of Physics: Conference Series*.
- 2
- 3 Selene Yue Xu. *Stock Price Forecasting Using Information from Yahoo Finance and Google Trend*. UC Berkeley.
- 4 Harry Markowitz. "Portfolio Selection: Efficient Diversification of Investments" (1952).
- 5 William F. Sharpe. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk" (1964).
- 6 George E.P. Box and Gwilym M. Jenkins.
- 7 Robert F. Engle. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation" (1982).
- 8
- 9 Luiz Henrique Barbosa Filho. Estimando a correlação usando um modelo GARCH no R. (2023). Disponível em: <https://analisemacro.com.br/data-science/dicas-de-rstats/estimando-a-correlacao-usando-um-modelo-garch-n>
- 10