

Arbitragem Estatística

Estudo em técnicas para análise de ativos financeiros

Livia Ohana da Rocha Carvalho

RA: 074178

Orientador: Lúcio Tunes

7 de dezembro de 2011

Agradecimentos

Este trabalho é produto de uma das oportunidades que a matemática aplicada me proporcionou que foi estagiar em um fundo de hedge ¹ em São Paulo (SP) chamado Principia Capital Management. Sendo assim, não posso deixar de agradecer aos meus gestores Marcello Paixão e Nelson Fernandes por me mostrarem que até as saídas de modelos matemáticos precisam ter um pouco de bom senso.

E ainda ao Professor Lúcio Tunes por me ajudar a fazer com que um pouco da experiência que eu adquiri na empresa faça parte da minha avaliação acadêmica. Por fim, mas não menos importante, um sincero agradecimento aos autores das teses e artigos que eu usei para escrever este trabalho pois sem eles o mesmo não seria possível.

Resumo

A Principia Capital Management é uma empresa de gestão de recursos com foco na maximização de retorno que utiliza-se intensivamente de ferramental quantitativo para estratégias de trading e para controle de risco [1]. Este trabalho será, basicamente, sobre estratégias financeiras usando o conceito de arbitragem estatística. Para isso precisamos reforçar conceitos matemáticos chave que explicam da validade de certos modelos.

Sumário

1	Introdução	2
2	Definições	2
2.1	Conceitos estatísticos e matemáticos	2
2.2	Alguns conceitos financeiros	3
3	Arbitragem Estatística	4
3.1	A estratégia	4
3.1.1	Trade de pares ou Long/Short	4
3.1.2	Modalidades de Long/Short	4
3.2	Backtest	5
3.2.1	Fatores do Backtest	5
4	Conclusões e perspectivas futuras	6

¹Podem ser definidos como fundos que adotam um número de estratégias que não podem ser adotadas por fundos tradicionais de investimento, mas isso não implica necessariamente se são mais ou menos arriscados. Embora o nome indique hedge (proteção, do inglês), isso não significa que todos os fundos desta categoria utilizem estratégias de hedge para proteger o desempenho de suas carteiras.

1 Introdução

A priori, para entendermos o conceito de arbitragem estatística precisamos entender em que consiste uma *arbitragem*. Uma arbitragem pura significa uma oportunidade de comprar um ativo e conseguir vendê-lo (ou um ativo “equivalente”) de maneira a obter um lucro sem risco [2]. É fácil inferir que este tipo de oportunidade não ocorre ou ocorre por pouquíssimo tempo no mercado.

Quando nos referimos a arbitragem estatística a principal diferença é que este tipo de lucro não é auferido sem risco, no entanto, busca-se um lucro esperado positivo e neste caso *esperado* significa que a oportunidade é lucrativa na média [2].

A estratégia para explorar esse tipo de evento, normalmente, se baseia na compra/venda de ativos toda vez que o diferencial de preços entre eles distorcer muito [10].

Os fundos que utilizam este tipo de modelo são chamados fundos quantitativos ou simplesmente fundos quant. A reta branca indica o CDI que é um índice de renda fixa, dando uma rápida olhada no gráfico da figura 1 podemos observar que 2011 não está sendo um bom ano para os fundos quantitativos já que todos eles estão rendendo menos que a renda fixa até setembro/2011 [3] (Dados: Bloomberg).

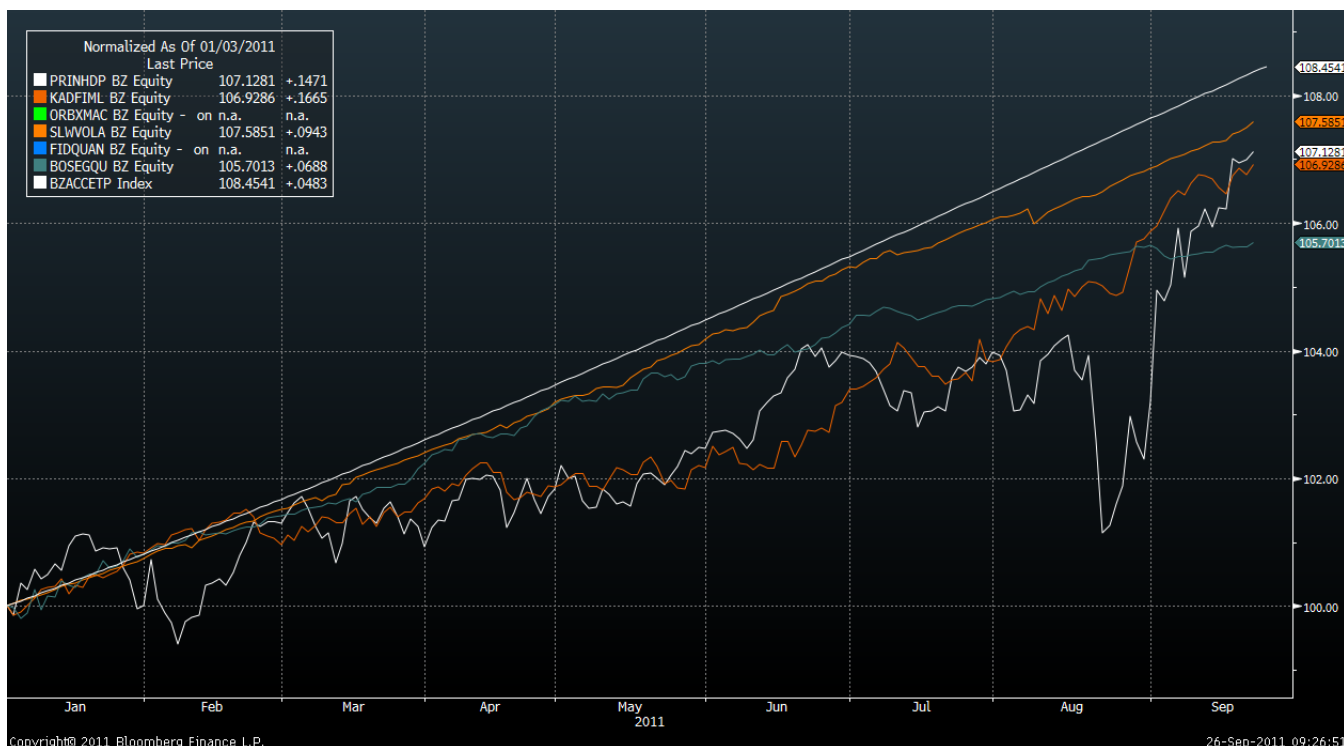


Figura 1: Gráfico de desempenho dos fundos quantitativos no Brasil de janeiro a setembro de 2011

2 Definições

2.1 Conceitos estatísticos e matemáticos

Definição 1 (Correlação). *Pode ser definido como a medida de dependencia entre duas quantidades [4]*

$$\rho_{X,Y} = \text{corr}(X,Y) = \text{cov}(X,Y) \div (\sigma_X \sigma_Y) = E[(X - \mu_X)(Y - \mu_Y)] \div (\sigma_X \sigma_Y) \quad (1)$$

Onde X, Y são variáveis aleatórias

μ_X, μ_Y são respectivamente as médias de X e Y

σ_X, σ_Y são respectivamente o desvio padrão de X e Y

Definição 2 (Estacionariedade ou processo estacionário). *É um processo estocástico, cuja distribuição de probabilidade conjunta não muda quando deslocado no tempo ou espaço. Consequentemente, parâmetros como a média e variância, se elas existem, também não se alteram ao longo do tempo ou posição [5].*

Definição 3 (Reversão à média). *Reversão à média (também conhecido como regressão à média) é o fenômeno estatístico onde temos que se uma variável é extrema em sua primeira medição, ela tenderá a ser mais próxima da média em uma segunda medição [6].*

Definição 4 (Séries temporais). *Uma série temporal é uma sequência de realizações (observações) de uma variável ao longo do tempo*

Uma série temporal pode ainda ser representada através de um componente determinístico e outro estocástico.

Definição 5 (Cointegração). *Se duas ou mais séries são individualmente integradas (no sentido de séries temporais), mas alguma combinação linear delas tem uma ordem inferior de integração, então a série se diz ser cointegradas [7]*

Aplicando estes conceitos ao mercado financeiro temos, por exemplo, que as séries de preços de ativos financeiros são, em geral, não-estacionárias, ié, média e variância variam no tempo. Outra característica dos preços de ativos é que eles em geral são integrados de ordem 1 (I(1)). Isso significa que os preços das ações podem subir indefinidamente (mas, obviamente, não podem cair abaixo de 0). Por outro lado, uma variável estacionária, integrada de ordem 0 (I(0)), apresenta uma característica interessante de sempre retornar à um valor central, ou médio. Dizemos que ela “reverte à média”. Os retornos diários de ações seguem, em geral, esse tipo de comportamento, apesar de não serem estritamente estacionários, já que suas médias (e principalmente suas volatilidades) variam no tempo. Para obter uma estratégia de arbitragem estatística bem-sucedida, é preciso trabalhar com uma variável que seja o mais próxima da estacionariedade possível. Afinal, queremos que haja convergência [2]. Certamente existem vários outros pontos a serem levados em consideração, mas a priori, usaremos definições mais simplistas.

Em alguns casos as séries temporais adquirem uma característica de retorno médio ao redor de zero, no entanto, a volatilidade (ou variabilidade) de dados muda com o tempo. De fato, os dados mostram que períodos de alta volatilidade tendem a ser agrupados. Um problema na análise deste tipo de dados financeiros é a previsão a volatilidade dos retornos futuros. Modelos como ARCH e GARCH modelos (Engle, 1982; Bollerslev, 1986) e modelos de volatilidade estocástica (Harvey, Ruiz e Shephard, 1994) foram desenvolvidos para lidar com esses problemas [13].

O critério de divergência será uma transformação dos preços das ações (série temporal que deverá ter uma característica estacionária ou pelo menos próxima de ser estacionária). Ou seja, para que a estratégia funcione é preciso um comportamento de reversão à média.

A partir dos trabalhos seminais de Markowitz (1959), Sharpe (1964), Lintner (1965) e Black (1972), a ferramenta estatística fundamental para otimização de portfólios é a análise da correlação dos retornos dos ativos. Os modelos de otimização para construção de portfólio, em sua maioria, focam em minimizar a variância de uma carteira de ativos, para um dado nível de retorno, com restrições adicionais a respeito dos investimentos permitidos, restrições quanto a venda a descoberto (short-sale) e os custos de transação associados ao rebalanceamento do portfólio [9].

A minimização da variância de uma carteira de ativos ajuda esta carteira a ser mais competitiva frente as outras pois esta variância pode ser entendida como o “risco” da carteira. Ou seja, quanto maior a variância maior a probabilidade de perdas e/ou ganhos. Carteiras com uma alta variância, em tese, precisam ter altos retornos médios para serem competitivas no mercado.

2.2 Alguns conceitos financeiros

Os conceitos abaixo são de suma importância no entendimento da estratégia de Long/Short que é uma das mais importantes estratégias em arbitragem estatística e está descrita em 3.1.

Definição 6 (Long - Comprado). *Uma posição long ou comprada é quando se possui uma determinada ação e se espera lucrar na alta, ié, com a valorização do papel.*

Definição 7 (Short - Vendido). *É possível entrar em uma posição short ou vendida, em geral, através do empréstimo de ações e esta operação faz sentido quando se acredita que o preço da ação vai cair (ganhar na*

baixa). A idéia é alugar as ações a uma pequena taxa (taxa esta que varia de ação para ação), vendê-las, recomprá-las quando o preço baixar, devolver o papel alugado. O lucro é calculado através da variação do preço da ação vezes o número de ações alugadas menos o custo do aluguel e taxas operacionais.

3 Arbitragem Estatística

O termo arbitragem estatística compreende estratégias nas quais buscam-se obter lucro a partir de anomalias no diferencial de preços entre pares de ativos [10], neste tipo de estratégia os parâmetros analisados são estritamente numéricos, logo não há uma análise qualitativa dos ativos envolvidos.

Abaixo faremos uma definição mais formal do termo, tal definição em um primeiro momento pode parecer obscura, mas será esclarecida no decorrer do trabalho.

Definição 8 (Arbitragem estatística). *Estratégia de alocar recursos em ativos “equivalentes” de maneira a obter um lucro esperado positivo [2].*

Para entendermos o conceito de ativos “equivalentes” precisamos entender a estratégia de trade de pares.

3.1 A estratégia

3.1.1 Trade de pares ou Long/Short

Definição 9 (Trades de pares). *É uma estratégia de especulação de curto prazo com uma lógica bastante simples [11]:*

- 1) *Encontre duas ações cujos preços variaram historicamente de maneira parecida (análise de correlação dos retornos dos ativos)*
- 2) *Esperar o spread (diferença de valores) entre os dois papéis aumentar*
- 3) *Ficar vendido (short) na ponta com preço mais alto*
- 4) *Ficar comprado (long) na ponta com preço mais baixo*
- 5) *Se a história se repete os preços convergem e a operação gerará lucro*

Se o mercado de ações fosse eficiente em todos os momentos os retornos ajustados neste tipo de operação não deveriam ser positivos.

Agora que o conceito de trade de pares foi esclarecido, talvez o conceito de arbitragem estatística fique mais claro. Podemos classificar os pares de ações como os ativos “equivalentes” da definição 8.

Resumindo, neste tipo de estratégia encontramos dois ativos que se comportem de maneira similar, esperamos que haja uma divergência entre eles e neste momento adotamos uma posição que acreditamos ser favorável pois se espera que, no futuro, haja uma convergência (i.e., a relação entre os ativos volta ao normal).

3.1.2 Modalidades de Long/Short

Abaixo encontram-se algumas definições de modalidade de trade de pares [8]. Tais modalidades ajudam a entender um pouco melhor como funciona a lógica da escolha dos pares em cada operação.

Definição 10 (intra-setorial). *O investidor compra e vende ações de empresas de um setor específico. É uma modalidade que tende a apresentar oscilações menores do que as demais, pois minimiza os riscos gerais, já que os impactos afetam o setor como um todo.*

Definição 11 (inter-setorial). *O investidor compra ações de um setor e vende ações de outro. É necessário um conhecimento profundo sobre os setores nessa modalidade, mais importante até do que o conhecimento das empresas propriamente dito.*

Definição 12 (carteira contra carteira). *O investidor compra uma carteira de ações diversas e vende outro grupo de ações, sem guardar relações entre setores. Uma variante dessa modalidade é a utilização de uma carteira de ações compradas e a venda de contratos futuros de índices de ações.*

Definição 13 (quantitativa). *Decisões são muito influenciadas por análises estatísticas.*

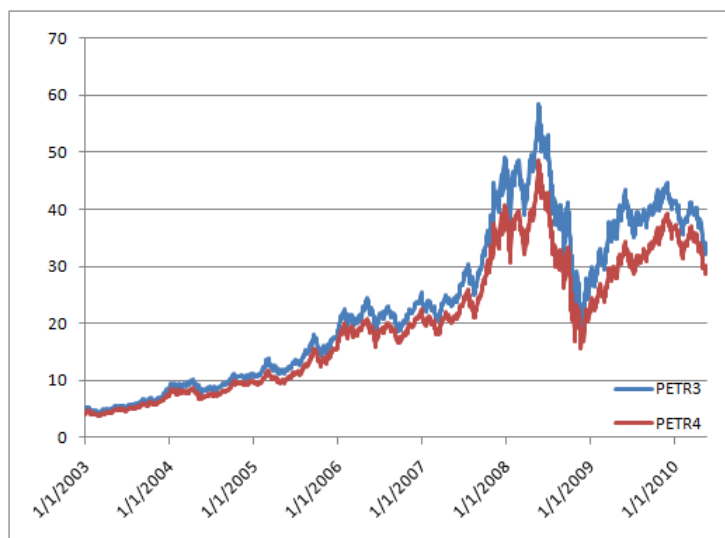


Figura 2: Gráfico que mostra o preço das ações PETR4 e PETR3

Na figura 2 temos um exemplo de ações que poderiam formar um par de Long/Short intra-setorial.

A figura 2 representa os preços das ações PETR3 é a ação ON (ordinária) e a PETR4 é a ação PN (preferencial), ambas da PETROBRAS e visto que são ações da mesma empresa, estão expostas aos mesmos riscos, e, como esperado, seus preços andam juntos. Uma estratégia de arbitragem estatística consiste em ao enxergar divergências no comportamento usual destas duas ações e fazer uma operação de Long/Short. Nota-se que identificar um critério de divergência é parte fundamental de toda a operação.

3.2 Backtest

Definição 14 (Backtest). *Simulação histórica dos trades que uma estratégia teria feito ao longo do tempo, usando em cada ponto do tempo apenas os dados disponíveis até lá.*

No exemplo de PETR3-PETR4 da figura 2, significaria percorrer a série temporal do spread entre as ações ao longo de um período pré-determinado, para um conjunto fixo de parâmetros, e anotar o desempenho de todos os trades gerados pelo modelo [2].

As estatísticas geradas pelo backtest ajudam a responder as seguintes perguntas:

- 1) Que outros pares seriam bons para esta estratégia?
- 2) Qual o retorno esperado dos trades por par?
- 3) Qual(is) o(s) risco(s) das posições?
- 4) A estratégia dá retorno positivo consistentemente? Se não, é possível identificar em que períodos ela não funciona?
- 5) Como muda o desempenho da estratégia com mudanças nos parâmetros?

3.2.1 Fatores do Backtest

Dados: Uma estratégia de investimento quant é um algoritmo que toma como valores de entrada dados (por exemplo, preços de ações) e retorna como saída uma ação (comprar, vender, zerar) a ser executada. É necessário checar a validade dos dados e o período em que se está fazendo o backtest e, se possível, usar um conjunto de ações que poderia ser realmente utilizado no período em questão.

Plataforma: É o programa de computador que vai realizar a simulação. Alguns exemplos comuns são Excel, Matlab, R, plataformas dedicadas e ainda programas em C, C++ e C#.

Modelo de trading e parâmetros: Neste momento é preciso traduzir o trade em termos analíticos, quantitativos e fugindo de ambiguidades. Poderíamos rodar um backtest simples para o exemplo da figura 2 usando 3 parâmetros como o tamanho da janela para calcular a média e desvio padrão do spread, o valor para entrar nos trades, e o valor para sair.

Custos: É preciso analisar os custos inerentes a qualquer estratégia de investimento. Os de operação incluem corretagem, emolumentos, aluguel de ações e ainda a diferença entre o preço observado e o que será realizado no trade (slippage). Com estes valores podemos calcular o impacto do custo sobre a estratégia.

4 Conclusões e perspectivas futuras

Posso dizer, que durante os meus meses de Principia vários conceitos tão exaustivamente cobrados durante todo o curso de matemática aplicada se tornaram, dentro do contexto, essencialmente simples. A principal estratégia do fundo é a estratégia de Long/Short e por isso decidi explorá-la melhor durante o trabalho.

O fundo em si possui uma base de dados considerável e é importante fazer um adendo sobre esses dados utilizados pela empresa para rodar os modelos, alguns são provenientes de programas internos e outros são dados coletados pela Economática².

Além disso, acompanhando um pouco das estratégias pude notar que apesar do fundo escolher as suas operações baseado nas simulações que, por sua vez, usam apenas dados quantitativos das empresas, é preciso ter uma visão holística de mercado para se escolher os resultados com maior chance de sucesso.

O objetivo deste trabalho foi introduzir o conceito arbitragem estatística como uma alternativa a análise de ativos. Certamente há ainda muito o que dizer e explicar sobre os modelos e critérios. Para os interessados nos modelos em si e em alguma implementação sugiro uma leitura mais atenta à referência 10.

Com relação às perspectivas futuras acredito que este tipo de abordagem, digamos mais analítica, tende a crescer mais no mercado brasileiro e já é amplamente utilizada não apenas por fundos de investimento, mas pelas maiores instituições financeiras mundiais.

Referências

- [1] <http://www.principiacm.com/>.
- [2] <http://drnickel.wordpress.com/>.
- [3] <http://www.bloomberg.com/>.
- [4] http://en.wikipedia.org/wiki/Correlation_and_dependence.
- [5] http://en.wikipedia.org/wiki/Stationary_process.
- [6] http://en.wikipedia.org/wiki/Regression_toward_the_mean.
- [7] <http://en.wikipedia.org/wiki/Cointegration>.
- [8] <http://www.chrinvestor.com>.
- [9] João Caldeira. Arbitragem estatística e estratégias de long-short, pairs trading, abordagem da cointegração aplicada a dados do mercado brasileiro. 2010.
- [10] Rafael Ferreira. Modelos de arbitragem estatística para o mercado brasileiro. 2011.
- [11] E.G. Gatev, W.N. Goetzmann, K.G. Rouwenhorst, and National Bureau of Economic Research. *Pairs trading: performance of a relative value arbitrage rule*. Working paper series. National Bureau of Economic Research, 1999.

²<http://www.economica.com/pt/>

- [12] A. Pole. *Statistical Arbitrage: Algorithmic Trading Insights and Techniques*. Wiley Finance. John Wiley & Sons, 2011.
- [13] R.H. Shumway and D.S. Stoffer. *Time series analysis and its applications*. Springer texts in statistics. Springer, 2000.