

4

Imunização utilizando análise de componentes principais

Até então, estudamos o que é, como funciona a imunização na teoria. Mostramos as vantagens de sua utilização e também suas limitações. Neste capítulo iremos aplicar um modelo de imunização com dados hipotéticos. O modelo escolhido para aplicação foi um desenvolvido por Barber e Cooper e publicado no *Journal of Portfolio Management* em 1996. O modelo tem como objetivo principal imunizar um fluxo de passivos. Para tal utiliza-se a análise de componentes principais para se determinar um número de direções básicas que possa direcionar, que possa prever os movimentos futuros na estrutura a termo. Inicialmente, utilizaremos apenas o primeiro componente como direcionador da curva para imunizarmos um fluxo de passivo. Em seguida, um número maior de componentes será utilizado.

4.1.

O modelo de imunização utilizando-se ACP

O modelo de imunização utilizando análise de componentes principais se divide em duas partes: um modelo mais simples, utilizando-se apenas um componente com diretor da curva e um outro modelo, mais geral, onde podemos utilizar mais componentes para explicarmos os movimentos na estrutura.

Para entendermos o modelo, precisamos também, definir como varia de um instante para outro a estrutura a termo das taxas de juros. O que será feito a seguir.

4.1.1.

Modelo de direção única

O modelo de direção única busca identificar uma única direção que possa antecipar mudanças na estrutura a termo. E com isso, adaptar a teoria de duração de Fisher-Weil (1971) para incluir esta direção (uma descrição detalhada deste tipo de duração se encontra no apêndice do trabalho). Vale lembrar, que a duração

de Fisher permite que a curva inicial seja de qualquer tipo, mas a partir daí, suas movimentações só podem ocorrer de forma paralela a esta curva inicial. Então, para iniciar a modelagem, precisamos definir como a curva de juros se move. Sabemos que taxas de juros para diferentes vencimentos são correlacionadas, ou seja, nos referimos a curva de juros em geral (as taxas subiram ou caíram). Neste primeiro caso, iremos supor que as taxas para diferentes vencimentos são perfeitamente correlacionadas. Esta suposição será relaxada no segundo modelo (de múltiplas direções).

Suponha que o conjunto inicial de taxas spot é conhecido e dado pela função $r_0(s)$, a qual determina uma taxa spot para cada data de vencimento s . Assumimos que estas taxas variam pelo valor de $u(s)h$, onde $u(s)$ é uma função conhecida da data de vencimento e h é uma variável aleatória. Então a taxa spot após um choque é dada por:

$$r(s) = r_0(s) + u(s)h \quad (4.1)$$

O preço de 1,00 prometido na data t_i é dado por $P_i(h) = \exp[-r(t_i)t_i]$. Suponha que n fluxos de ativos ou passivos serão recebidos ou pagos nas datas $t_1, \dots, t_n$. O valor presente do fluxo de caixa $C_1, \dots, C_n$ é dado por:

$$S(h) = \sum_{i=1}^n P_i(h)C_i \quad (4.2)$$

Podemos aproximar uma mudança no valor do excesso relativa a uma pequena variação na variável aleatória h por

$$\frac{dS(h)}{dh} = \sum_i \frac{dP_i(h)}{dh} C_i \quad (4.3)$$

Aplicando a regra da cadeia, temos:

$$\frac{dP_i(h)}{dh} = \frac{dP}{dr} \frac{dr}{dh} = -t_i \exp[-r(t_i)t_i] u(t_i) = -t_i P(h) u(t_i) \quad (4.4)$$

Substituindo na equação (4.3):

$$\frac{dS(h)}{dh} = \sum_i -t_i P(h) u(t_i) C_i \quad (4.5)$$

Escrevendo na forma de diferenças finitas:

$$\begin{aligned}
\Delta S &\approx \Delta h \sum_{i=1}^n P'_i(0) C_i \\
&= -\Delta h \sum_{i=1}^n P_i(0) C_i u(t_i) t_i
\end{aligned}
\tag{4.6}$$

A substituição de h por zero foi feita para indicar que estamos no tempo zero.

A imunização tradicional requer a escolha de um fluxo de caixa de ativos, dado um fluxo de caixa de passivos, de forma que a aproximação linear de $\Delta S = 0$. Para o modelo de mudança paralela, definimos $u(s)$ igual a constante k e então temos

$$\frac{\Delta S}{S} \approx -\Delta h k D_{FW} \tag{4.7}$$

Onde

$$D_{FW} = \frac{1}{S(0)} \sum_{i=1}^n P_i(0) C_i t_i \tag{4.8}$$

é a duração de Fisher-Weil (seção 11.2.7) para o valor de excesso. O modelo de imunização para mudanças paralelas requer escolher fluxos de caixa de modo que $D_{FW} = 0$.

No modelo de variação paralela, o formato da curva não muda ao longo do tempo e a volatilidade é a mesma para todas as datas de vencimento. Neste modelo de direção única, dependendo da especificação de $u(s)$, o formato da curva irá variar durante o tempo e a volatilidade será uma função do vencimento.

4.1.2. Modelo de múltiplas direções

Definimos acima, um modelo de única direção. Agora, como dito, iremos relaxar a restrição de que as taxas são perfeitamente correlacionadas. Neste modelo, as taxas para diferentes vencimentos não são perfeitamente correlacionadas. Para isso, precisamos adicionar outros fatores ao modelo.

Novamente, assumimos que o conjunto inicial de taxas spot é dado por $r_0(s)$, só que agora, a curva após um choque é definida por uma série de fatores. Definimos então $x(s)$ como a mudança na taxa spot:

$$x(s) = r(s) - r_0(s) \tag{4.9}$$

As taxas spot agora assumimos evoluírem de acordo com o seguinte modelo multifator:

$$x(s) = \sum_{k=1}^K u_k(s) h_k \quad (4.10)$$

Suponha que N fluxos de caixa são esperados serem pagos ou recebidos nas datas $t_1, \dots, t_N$. Podemos expressar $x(s)$ e $u_1(s), \dots, u_k(s)$ como vetores coluna de dimensão N.

Seja $X = [x(t_1), \dots, x(t_N)]^T$ e $U_1 = [u_1(t_1), \dots, u_1(t_N)]^T$ assim por diante.

Reescrevendo a equação de evolução das taxas em notação vetorial,

$$X = \sum_{k=1}^K h_k U_k \quad (4.11)$$

$$u_i u_j = 0 \quad \begin{matrix} i = 1, \dots, k \\ j = 1, \dots, k \end{matrix} \quad i \neq j \quad (4.12)$$

$$|u_i| = 1 \quad i = 1, \dots, k \quad (4.13)$$

Sem perda de generalidade, podemos assumir que $U_1, \dots, U_k$ são ortonormais. Cada vetor ortonormal determina uma direção fundamental na qual as taxas spot podem mudar. Qualquer vetor de mudança X pode ser expresso como uma combinação linear do K vetores fundamentais.

Para um portfolio ser imunizado, ele deve ser propriamente protegido contra mudanças em cada direção fundamental. Para um dado fluxo de passivos, requer escolher um fluxo de caixa de ativos de forma que as K condições sejam satisfeitas:

$$\sum_{i=1}^n P_i(0) C_i u_k(t_i) t_i = 0 \quad (4.14)$$

$$k = 1, \dots, K$$

Ao contrário do modelo de direção única, o modelo de múltipla direção permite que o formato da curva de juros varie de varias formas diferentes. O problema é escolher o conjunto fundamental das direções dos vetores.

4.2. Análise de Componentes Principais (ACP)

A análise de componentes principais (principal component analysis - PCA) tem como objetivo principal a obtenção de um pequeno número de combinações lineares (componentes principais) de um conjunto de variáveis, que retenham o máximo possível da informação contida nas variáveis originais. O que se observa em vários estudos é que, freqüentemente, um pequeno número de componentes pode ser usado, no lugar das variáveis originais, nas análises de regressões. O trabalho mais conhecido neste assunto é o de Litterman e Scheinkman. Neste artigo, de maneira exploratória, os autores identificam 3 componentes capazes de explicar algo em torno de 98% da variabilidade das taxas implícitas a papéis de vários vencimentos no mercado americano. Além disso, Litterman e Scheinkman interpretam estes 3 componentes como sendo responsáveis, em ordem de importância, pelos seguintes movimentos da curva de juros: variação de nível (mudanças paralelas), mudança de inclinação e mudança na curvatura.

Os componentes são extraídos na ordem do mais explicativo para o menos explicativo e sabemos que na teoria o número de componentes é sempre igual ao número de variáveis. Entretanto, alguns poucos componentes são responsáveis por grande parte da explicação total.

Para esta aplicação de imunização, nosso objetivo é extrair do histórico de variações da curva de juros uma pequena família de direções fundamentais que descreva aproximadamente o histórico de variações das taxas no período selecionado. E que sirvam como um modelo de previsão da movimentação da curva a partir daí.

Veremos a seguir, matematicamente, o que significa este tipo de análise.

Seja X_t um vetor linha, de dimensão $(1 \times k)$, de uma observação no tempo t , ou seja:

$$X_t = [x_{t1} \quad x_{t2} \quad \cdots \quad x_{tk}] \quad (4.15)$$

Supondo que fizemos n observações, podemos montar uma matriz X de n observações em k variáveis.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad (4.16)$$

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix} \quad (4.17)$$

onde as observações são expressas na forma de desvio da média da amostra, na qual estamos interessados em estudar a variação nos dados.

A natureza dos componentes principais pode ser aproximada de varias maneiras. Uma delas é perguntar quanta independência realmente existe no conjunto de k variáveis. Ou ainda, mais explicitamente consideramos a transformação dos elementos da matriz X em um novo conjunto de variáveis que seja dois a dois descorrelacionado e que o primeiro tenha a maior variância possível, o segundo a segunda maior variância possível e assim por diante.

Para tal, definiremos z_{1t} como uma combinação linear dos elementos da t -ésima linha da matriz X para $t = 1, \dots, n$.

$$z_{1t} = a_{11}x_{t1} + a_{21}x_{t2} \dots + a_{k1}x_{tk}$$

Escrevendo em forma matricial temos:

$$z_{1t} = \begin{bmatrix} x_{t1} & x_{t2} & \cdots & x_{tk} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{bmatrix} = X_t a_1 \quad (4.18)$$

Ou ainda, definindo z_1 como a matriz de todos z_{1t} :

$$z_1 = X a_1 \quad (4.19)$$

z_1 é um vetor $(n \times 1)$ e a_1 é um vetor $(k \times 1)$.

Obs:

$$[n \times 1] = [n \times k] \cdot [k \times 1] \quad (4.20)$$

A soma dos quadrados de z_1 em forma matricial é:

$$z_1' z_1 = a_1' X' X a_1 \quad (4.21)$$

Como nosso objetivo é maximizar a variância, devemos escolher a_1 de forma a maximizar $z_1' z_1$. Claramente temos que inserir algumas restrições ao

problema de maximização, caso contrário, $z_1'z_1$ poderia ser feito infinitamente grande. Para isto, iremos normalizar o vetor a_1 . Logo:

$$a_1'a_1 = 1 \quad (4.22)$$

Nosso problema torna-se então maximizar $z_1'z_1$ sujeito a $a_1'a_1 = 1$. Para tal, utilizamos o conceito do multiplicador de Lagrange. Nossa função fica:

$$\phi = a_1'X'Xa_1 - \lambda_1(a_1'a_1 - 1) \quad (4.23)$$

onde λ_1 é o multiplicador de Lagrange.

Derivando a expressão acima (4.23) e aplicando-se as condições de primeira ordem obtemos:

$$\frac{\partial \phi}{\partial a_1} = 2X'Xa_1 - 2\lambda_1(a_1) = 0 \quad (4.24)$$

Arrumando a expressão (4.24), chegamos que

$$(X'X)a_1 = \lambda_1(a_1) \quad (4.25)$$

Esta é uma equação conhecida da teoria de autovalores e autovetores. Logo, podemos concluir que a_1 é um autovetor da matriz $(X'X)$ correspondente ao autovalor λ_1 . Substituindo em $z_1'z_1$ temos

$$z_1'z_1 = \lambda_1 a_1'a_1 = \lambda_1 \quad (4.26)$$

Então, devemos escolher λ_1 como sendo o maior autovalor da matriz $(X'X)$.

A matriz $(X'X)$, na ausência de colinearidade perfeita, será positiva definida e então terá autovalores positivos (veja a seção 11.3.5 – expressão (A.75)). O primeiro componente principal de X será então z_1 .

O segundo passo é definir $z_2 = Xa_2$. Logo,

$$z_2'z_2 = a_2'X'Xa_2 \quad (4.27)$$

Só que neste caso queremos escolher a_2 de forma a maximizar $a_2'X'Xa_2$ sujeito as restrições $a_2'a_2 = 1$ e $a_1'a_2 = 0$. A razão para a segunda condição é que z_2 seja não correlacionada com z_1 . A covariância entre elas é dada por

$$\begin{aligned} a_1'X'Xa_2 &= \lambda_1 a_1'a_2 \\ &= 0 \quad \text{se} \quad a_1'a_2 = 0 \end{aligned} \quad (4.28)$$

Montando-se novamente a equação de maximização, neste caso com dois multiplicadores de Lagrange (duas restrições):

$$\phi = a_2' X' X a_2 - \lambda_2 (a_2' a_2 - 1) - \mu (a_1' a_2) \quad (4.29)$$

onde λ_2 e μ são os multiplicadores de Lagrange.

Derivando e aplicando as condições de primeira ordem temos:

$$\frac{\partial \phi}{\partial a_2} = 2X'Xa_2 - 2\lambda_2 a_2 - \mu a_1 = 0 \quad (4.30)$$

Premultiplicando-se a equação por a_1' e utilizando-se das definições acima obtemos:

$$2a_1' X' X a_2 - \mu = 0 \quad (4.31)$$

Sabemos da primeira maximização que

$$(X'X)a_1 = \lambda_1 a_1 \quad (4.32)$$

Então, premultiplicando por a_2' temos:

$$a_2' (X'X)a_1 = \lambda_1 a_2' a_1 = 0 \quad (4.33)$$

Donde se conclui que

$$\mu = 0 \quad (4.34)$$

O que nos conduz a seguinte expressão:

$$X'Xa_2 = \lambda_2 a_2 \quad (4.35)$$

E, exatamente da mesma maneira feita acima, devemos escolher λ_2 como sendo o segundo maior autovalor da matriz $(X'X)$.

Podemos continuar este processo para cada um dos k autovalores de $(X'X)$ e agrupar os vetores resultantes em uma matriz ortogonal que chamaremos de A (matriz de autovetores de $(X'X)$).

$$A = [a_1 \quad a_2 \quad \cdots \quad a_k] \quad (4.36)$$

Logo, os k componentes principais de X podem ser agrupados em uma matriz $(n \times k)$ chamada de Z :

$$Z = XA \quad (4.37)$$

Obs:

$$[n \times k] = [n \times k] \cdot [k \times k] \quad (4.38)$$

Além disso

$$Z'Z = A'X'XA = \Lambda \quad (4.39)$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \lambda_k \end{bmatrix} = Z'Z \quad (4.40)$$

O que mostra que os componentes principais são realmente não correlacionados dois a dois e que suas variâncias são dadas por:

$$z_i' z_i = \lambda_i \quad i = 1, \dots, k \quad (4.41)$$

Caso o posto (*rank*) da matriz X seja $r < k$, $k - r$ autovalores serão zero e as variações em X podem ser completamente expressas em termos das r variáveis independentes. Além disso, mesmo se X tiver *full column rank* (rank da matriz é igual ao número de colunas) é, alguns dos seus autovalores podem estar bem próximos de zero de forma que um pequeno número de componentes principais respondam por uma grande parte da variância de X . A variação total nos X 's é dada por:

$$\sum_t x_{t1}^2 + \sum_t x_{t2}^2 + \dots + \sum_t x_{tk}^2 = tr(X'X) \quad (4.42)$$

Da propriedade traço de uma matriz, sabemos que:

$$tr(AB) = tr(BA) \quad (4.43)$$

Usando-se esta propriedade e o fato de que A é uma matriz ortogonal:

$$AA' = I \quad (4.44)$$

Podemos dizer que:

$$tr(A'X'XA) = tr(X'XAA') = tr(X'X) \quad (4.45)$$

Podemos ainda escrever o traço de $(X'X)$ na forma de somatório e usando a propriedade acima podemos mostrar que:

$$\sum_{i=1}^k \sum_{t=1}^n x_{it}^2 = tr(X'X) = \sum_{i=1}^k \lambda_i = z_1' z_1 + \dots + z_k' z_k \quad (4.46)$$

Nosso último passo na análise de componentes principais é saber quanto cada componente tem de poder de explicação da variância total dos dados iniciais. Para isso, utilizamos os autovalores. Para encontrarmos uma medida, dividiremos cada autovalor encontrado pelo somatório de todos autovalores da matriz $(X'X)$. Isto nós dará uma relação percentual do poder de explicação de cada componente.

$$\frac{\lambda_1}{\sum \lambda}, \frac{\lambda_2}{\sum \lambda}, \dots, \frac{\lambda_k}{\sum \lambda} \quad (4.47)$$

São estas frações que irão nos dizer quantos componentes são necessários para se ter uma boa estimativa, um bom poder de explicação. Inicialmente, deve-se fixar um percentual (usualmente utiliza-se algo entre 97%, 98%) no qual, só se utilizará os componentes até que seu poder de explicação acumulado atinja o valor percentual predeterminado.