

# 1

## Introdução

No Brasil, o grande aumento na adesão a planos de previdência privada nos últimos tempos implicou em maiores preocupações de bancos e seguradoras em fazer investimentos visando garantir o pagamento de benefícios no longo prazo. No entanto, o cenário da economia que hoje em dia é muito favorável pode não o ser daqui a alguns anos. Uma boa previsão do comportamento das taxas de juros é uma tarefa fundamental. Neste contexto, conseqüentemente, especificar e estimar modelos que forneçam previsões confiáveis e que acompanhem a curva de retornos com eficiência se torna cada vez mais importantes. Nesse sentido, modelos tradicionalmente propostos incluem Vetor Auto-regressivo (VAR) e Modelos de Correção de Erros (ECM). Nessa dissertação, procura-se desenvolver um novo modelo baseado na metodologia de aprendizagem estatística supervisionada.

A teoria estatística de aprendizagem busca extrair tendências e padrões importantes dos dados e entender o que os mesmos querem dizer. Conforme colocado por Vapnik (2000)(30), pode-se descrever o problema geral de aprendizagem por exemplos através de três componentes, conforme ilustrado na figura 1.1:

- Um gerador (G) de vetores aleatórios  $\mathbf{x} \in \mathbb{R}^n$ , obtidos independentemente a partir de uma distribuição de probabilidade fixa e desconhecida  $F(\mathbf{x})$ .
- Um supervisor (S) que retorna um valor de saída  $y$  para cada vetor de entrada  $\mathbf{x}$ , de acordo com uma função de distribuição condicional  $F(y | \mathbf{x})$ , também fixa e desconhecida.
- Uma máquina de aprendizagem (LM) capaz de implementar um conjunto de funções  $f(\mathbf{x}, \alpha)$ ,  $\alpha \in A$ , onde  $A$  é um conjunto de parâmetros.

Dado um conjunto de funções  $f(\mathbf{x}, \alpha)$ ,  $\alpha \in A$ , o problema de aprendizagem é escolher aquela que melhor se aproxima da resposta do supervisor. A seleção da função desejada é baseada em um conjunto de treinamento composto de  $l$  observações  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)$ , independentes e identicamente distribuídas (iid) obtidas de acordo com  $F(\mathbf{x}, y) = F(\mathbf{x})F(y | \mathbf{x})$ .

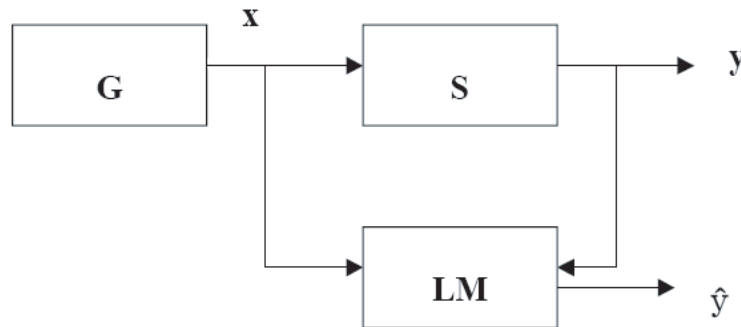


Figura 1.1: Modelo de aprendizagem supervisionada a partir de exemplos (Vapnik, 2000(30)).

Por sua vez, a aprendizagem supervisionada procura aprender uma função que forneça a saída desejada a partir de um conjunto de entradas usando exemplos (pares de entrada e saída desejada). Quando existe uma função dos dados de entrada para os dados de saída tal função é chamada de função alvo e a estimativa de tal função é escolhida dentre um conjunto de funções candidatas, conhecido como um conjunto de hipóteses ou espaço de hipóteses. Em aprendizagem supervisionada, também se deve ter boas propriedades de generalização para dados não utilizados no treinamento, chamados dados de teste.

Um método de aprendizagem supervisionada para classificação originalmente desenvolvido por Vapnik e colaboradores, conhecido como Máquina de Suporte Vetorial (SVM) utiliza um espaço de hipóteses de funções lineares em um espaço característico de alta dimensionalidade e tem sido aplicado com sucesso em tarefas, como por exemplo, OCR (reconhecimento de caráter ótico) (30, 20), classificação de tipos de rocha em geologia e geofísica (12, 13, 31) e, recentemente, em regressão (19, 10) e até alcançando êxito na previsão de séries temporais (25, 11). Por essa razão, estudamos Máquinas de Suporte Vetorial com ênfase na sua extensão para Regressão (SVR) visando sua aplicação em previsão de séries temporais.

Investiga-se ainda a utilização do SVR na previsão de retornos de títulos de renda fixa com diversas maturidades. Nessa formulação, além dos retornos, a informação contida em algumas variáveis macro-econômicas é explorada a fim de melhor prever o desenvolvimento da estrutura a termo. A idéia de utilizar variáveis macro-econômicas e os retornos foi desenvolvida para a economia brasileira (9) e verificou-se que, a informação contida nessas variáveis pode

melhorar o desempenho da previsão no longo prazo. Assim, formula-se o modelo SVR e os modelos tradicionalmente propostos (VAR e ECM), incluindo as observações passadas dos retornos e de algumas variáveis macroeconômicas já que o objetivo é obter um bom desempenho preditivo no longo prazo, que é de fundamental importância para fins atuariais.

## **1.1 Motivação**

Empresas de fundos de pensão têm como objetivo conseguir o máximo de lucro dos prêmios de seguro de pensão, evitando níveis de risco significantes, a fim de garantir que seus compromissos com o pagamento dos benefícios futuros sejam cumpridos. Ademais, dependendo da situação da instituição e das preferências de seus participantes, podem ocorrer diferentes estratégias de investimentos. Portanto, modelos que estudem o comportamento dos ativos e passivos ao longo do tempo (ALM) e aplicados a fundos de pensão estão diretamente associados aos riscos de longo prazo, os quais estão relacionados aos riscos de mercado, como o risco associado à variação da taxa de juros.

A Estrutura a Termo da Taxa de Juros (ETJT) é um indicador de grande importância na economia uma vez que, conceitualmente, representa a relação, em determinado instante, entre o prazo para o vencimento e a taxa de retorno dos títulos oriundos de uma mesma classe de risco, desempenhando um papel central na determinação de ativos (ou passivos) que possam ser representados por um fluxo de caixa.

Modelos econométricos realistas são naturalmente necessários, os quais devem levar em conta a evolução das variáveis macro-econômicas principais para prever as trajetórias dos retornos de investimentos e fluxo de responsabilidades. Na verdade, uma tentativa de resolver problemas de casamento de ativos e passivos (ALM) deve levar em consideração estimativas da estrutura a termo da taxa de juros, motivo pelo qual se faz obrigatório a existência de modelos que sejam confiáveis e bons preditores. Os modelos de Vetores Auto-regressivos (VAR) e Modelos de Correção de Erros (ECM) são as ferramentas tipicamente usadas para tais objetivos.

Esses modelos são freqüentemente utilizados para descrever o comportamento de um pequeno conjunto representativo de ativos composto, por exemplo, por títulos de renda fixa de curto e longo prazos, ações, móveis e indicadores de crescimento de salários e preços (cf. Boender, van Aalst e Heemskerk em (1) e o trabalho de Dert (8)). Assim, os modelos fornecem uma primeira avaliação dos retornos de investimentos desde que a maior parte da carteira típica de um fundo de previdência possa ser classificada nas categorias citadas.

Além disso, os modelos também fornecem uma primeira avaliação da evolução das contribuições e dos passivos, que estão relacionados ao crescimento dos salários e dos preços.

Esses modelos, no entanto, incorrem em algumas desvantagens relevantes:

- Não conseguem produzir uma descrição confiável da trajetória futura do valor de mercado das carteiras devido ao número limitado de categorias de ativos, dado que, na prática, essas carteiras não se limitam a ativos que acompanham os índices que aparecem nos modelos mais convencionais;
- Não fornecem os fatores de desconto necessários para calcular o preço de mercado dos diversos ativos de renda fixa disponíveis, sejam eles " *pure discount bonds*" ou " *coupon bearing bonds*"; e
- Não conseguem tratar adequadamente a questão do risco de reinvestimento, que necessariamente requer previsões para a rentabilidade de títulos de diversas maturidades.

Exercícios de previsão são geralmente realizados baseados nas hipóteses de que as variáveis relevantes (ou a combinação linear delas) são estacionárias e/ou combinadas por relações lineares. Estas restrições, entretanto, não são consistentes com as características de muitas séries temporais financeiras e econômicas, as quais não possuem médias e variâncias constantes. Um bom exemplo seriam as taxas de juros, nosso principal foco, as quais exibem períodos sucessivos de muita volatilidade e, pelo menos em princípio poderiam estar relacionadas umas com as outras e com outras variáveis macro-econômicas de diferentes maneiras dependendo do nível de volatilidade. Devido a essas características, é proposto aplicar um método de aprendizagem supervisionada chamado Máquina de Suporte Vetorial para Regressão (SVR) como uma técnica de modelagem, visando obter uma estimação mais precisa da curva de retornos, que é uma condição necessária para recuperação dos fatores de desconto e análise do risco de reinvestimento, desde que tal técnica facilmente acomoda relações não-lineares entre as variáveis relevantes.

## 1.2

### Objetivos

Máquinas de Suporte Vetorial para Regressão são uma nova metodologia de aprendizagem supervisionada não-paramétrica para estimação de regressão. Dentre as características principais da técnica, destacam-se o bom desempenho de generalização e o fato de que treinar SVR é equivalente a resolver um problema de programação quadrática cuja solução é sempre única e global, não existindo portanto problemas de existência de mínimos locais. Por fim,

consegue-se ainda uma representação esparsa da solução, pois a solução do SVR depende apenas de um subconjunto dos dados de entrada referidos como vetores de suporte.

Por todos esses motivos, examina-se a viabilidade de sua aplicação na previsão de retornos de investimentos em títulos de renda fixa de várias maturidades comparando com os modelos VAR e ECM, utilizando diversas métricas.

### 1.3

#### Estrutura da dissertação

No capítulo seguinte, algumas definições e teoremas da teoria de otimização que serão importantes para melhor entendimento da teoria SVR serão apresentados.

No capítulo 3 uma introdução a teoria de aprendizagem supervisionada é feita, bem como as vantagens e desvantagens de tal metodologia. A teoria de Máquina de Suporte Vetorial é estudada e uma análise em detalhes de sua extensão para regressão é realizada. Defini-se o princípio da minimização do risco empírico e sua importância para a solução do problema de otimização envolvido, o qual é descrito em detalhes, explicando-se ainda as vantagens em utilizar a representação dual do problema e sua solução. Algumas funções de custo possíveis são apresentadas e a definição da função de custo implementada, a saber, a função de custo  $\varepsilon$ -insensível é estudada. Considera-se a necessidade de introduzir as funções núcleo para incorporar as relações não-lineares dos dados. As propriedades das funções núcleo são analisadas bem como sua principal caracterização pelo Teorema de Mercer. Estuda-se ainda o uso do *software* R, que foi utilizado para a implementação do SVR.

No capítulo 4, é feita uma descrição dos dados utilizados e no capítulo 5, os modelos SVR com as medidas de erro utilizadas são apresentados, bem como a divisão dos dados de acordo com a teoria de aprendizagem.

No capítulo 6, os resultados e tabelas com as métricas usadas para comparar os métodos estudados são mostrados. Por fim, no capítulo 7, conclui-se o trabalho, alguns trabalhos futuros são discutidos e os gráficos comparativos entre os métodos tradicionais e os modelos SVR são exibidos.