

SCC 0270 – Redes Neurais e Aprendizado Profundo

Máquinas de Vetores de Suporte

Support Vector Machines – SVM

Prof. Bruno S. Faíçal

e-mail: bsfaical@alumni.usp.br

PAE: José Pedro Ribeiro Belo

e-mail: josepedroribeirobelo@gmail.com

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

Conteúdo visto

- Objetivo almejado no treinamento
- Métodos de avaliação
- Medidas de desempenho
 - Classificação
 - Regressão

Sumário

- 1 Recapitulação
- 2 **Introdução**
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

Introdução

- Máquinas de vetores de suporte (*support vector machines* – *SVM*) são fundamentadas na Teoria do Aprendizado Estatístico

Introdução

- Máquinas de vetores de suporte (*support vector machines* – *SVM*) são fundamentadas na Teoria do Aprendizado Estatístico
- De maneira geral e resumida, esta técnica busca pelo classificador ou função de regressão (dependendo da natureza do problema) **ótima**

Introdução

- Máquinas de vetores de suporte (*support vector machines* – *SVM*) são fundamentadas na Teoria do Aprendizado Estatístico
- De maneira geral e resumida, esta técnica busca pelo classificador ou função de regressão (dependendo da natureza do problema) **ótima**
- Esse objetivo é alcançado ajustando os hiperplanos e/ou funções em posições “ótimas”

Introdução

- Máquinas de vetores de suporte (*support vector machines* – *SVM*) são fundamentadas na Teoria do Aprendizado Estatístico
- De maneira geral e resumida, esta técnica busca pelo classificador ou função de regressão (dependendo da natureza do problema) **ótima**
- Esse objetivo é alcançado ajustando os hiperplanos e/ou funções em posições “ótimas”
- Para isso, a SVM utiliza uma abordagem chamada *vetor de suporte*

Introdução

- Máquinas de vetores de suporte (*support vector machines* – *SVM*) são fundamentadas na Teoria do Aprendizado Estatístico
- De maneira geral e resumida, esta técnica busca pelo classificador ou função de regressão (dependendo da natureza do problema) **ótima**
- Esse objetivo é alcançado ajustando os hiperplanos e/ou funções em posições “ótimas”
- Para isso, a SVM utiliza uma abordagem chamada *vetor de suporte*

Objetivo dessa aula

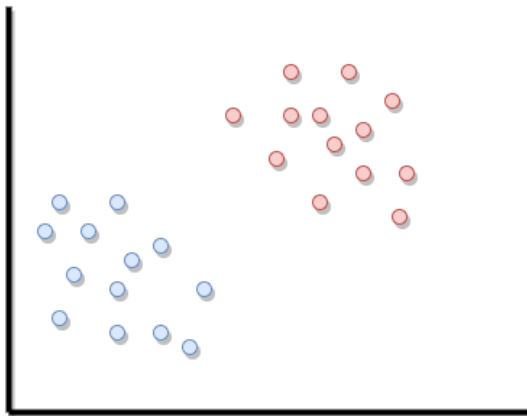
Entender o funcionamento geral da técnica SVM (em problemas de classificação e regressão) sem aprofundar-se nos conceitos matemáticos envolvidos. Desta forma, espera-se que após a aula, a execução da SVM seja compreensível e possa ser utilizada de maneira adequada para a resolução de diferentes problemas.

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

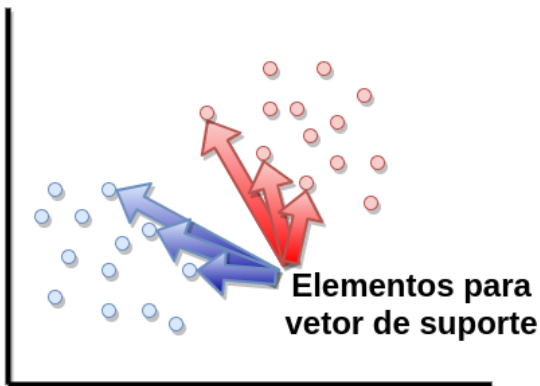
Vetor de Suporte

- Considere um problema de classificação binária (duas classes)



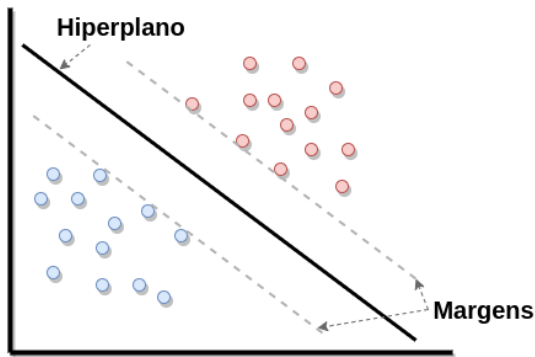
Vetor de Suporte

- Os elementos mais próximos da fronteira entre os grupos são selecionados para compor o vetor de suporte



Vetor de Suporte

- O vetor de suporte é utilizado para ajustar a posição e inclinação do hiperplano



Vetor de Suporte

Vetores de suporte

Portanto, podemos ver o elemento “vetor de suporte” como estruturas compostas por exemplos que auxiliam no ajuste dos hiperplanos para alcançarem posições *ótimas*.

Vetor de Suporte

Porque utilizar vetores de suporte?

Um hiperplano que adéque suas fronteiras aos elementos de borda das distribuições que pretende separar (i.e., posicione-se com a mesma distância entre elas), tende a possibilitar maior generalização do conhecimento fornecido pelas distribuições. Na prática, esse ajuste resulta em menos erro para exemplos não conhecidos.

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

SVM linear

- Inicialmente proposto para a resolução de problemas linearmente separáveis
- Esta é a formulação mais simples da SVM
- Proporciona o melhor “hiperplano” possível com base nos exemplos de treinamento
- Realiza divisões lineares com base na equação: $h(x) = w.x + b$
- Permitindo a divisão do espaço em duas regiões, tais como:

$$g(x) = \begin{cases} +1 & \text{se } w.x + b \geq 0 \\ -1 & \text{se } w.x + b \leq 0 \end{cases} \quad (1)$$

SVM linear

- Pode ser dividida em dois grupos:
 - Margens rígidas
 - Margens suaves
- A principal diferença está em como os vetores de suporte são usados
- Estes grupos também existem em SVM não-lineares

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

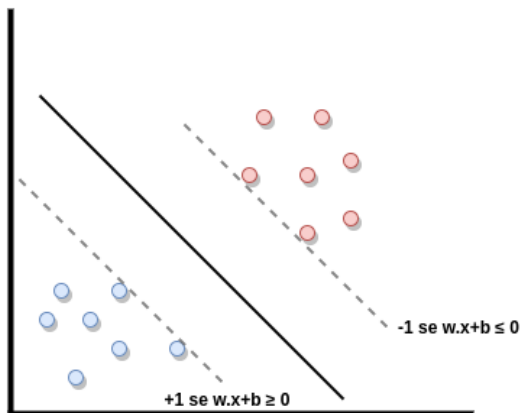
SVM linear – margens rígidas

- O elementos mais próximos da margem definem a sua posição
- Permanecem do lado correto referente a classe que pertence
- Impõe-se que não existam exemplos de treinamento entre as margens e o hiperplano de divisão de classes
- A definição de hiperplanos canônicos (margens) passam a ser tratadas como um problema de otimização dos pesos, tal como:

$$\text{Minimizar } \frac{1}{2} ||w||^2$$

SVM linear – margens rígidas

Exemplo da definição de hiperplanos com margens rígidas



Sumário

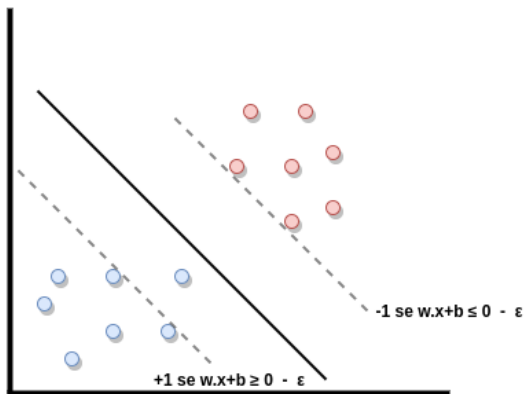
- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear**
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

SVM linear – margens suaves

- Em ambientes reais é difícil encontrar aplicações que possuem a separação dos exemplos de maneira linear de forma rígida
- Isso por ocorrer por diversos motivos, tais como: a própria natureza do problema ou a existência de *outliers*
- Assim, SVM com margens suaves permite que existam exemplos entre a margem e o hiperplano de divisão
 - Também a presença de alguns erros, quando o exemplo da classe $+1$ está localizado na região da classe -1 (definido pelo hiperplano de divisão)
- Isso é feito através de uma variável de folga, a qual relaxa a restrição anterior (das margens rígidas)

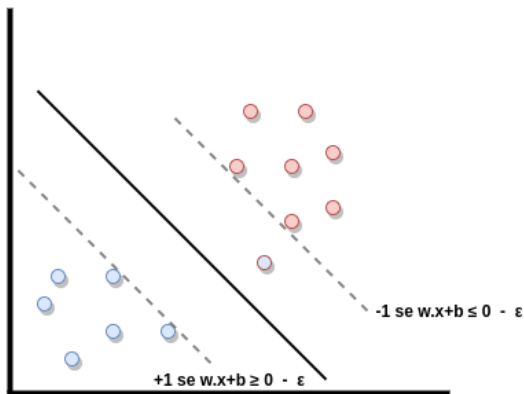
SVM linear – margens suaves

Exemplo da definição de hiperplanos com margens suaves – **sem** erro na classificação



SVM linear – margens suaves

Exemplo da definição de hiperplanos com margens suaves – **com** erro na classificação



SVM linear – margens suaves

- A definição de hiperplanos canônicos (margens) tem sua equação do seu problema atualizado para:

$$\text{Minimizar } \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^n \varepsilon_i \right)$$

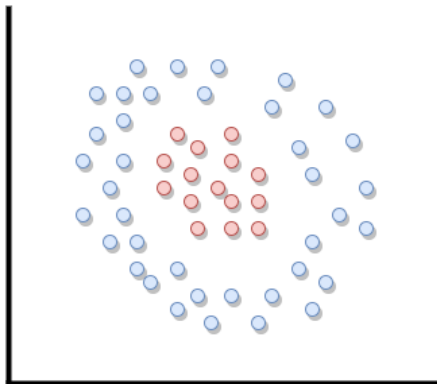
onde, C é o termo de regularização que impõe um peso à minimização dos erros no conjunto de treinamento

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

SVM não-linear

- Em alguns casos pode não ser possível separar satisfatoriamente os exemplos de maneira linear
- Por exemplo:



SVM não-linear

- Problemas não lineares são abordados utilizando *kernels*
- Um *kernel* é responsável em mapear um conjunto de treinamento de seu espaço original (entradas) para um novo espaço com maior quantidade de dimensões (espaço de características – *feature space*)

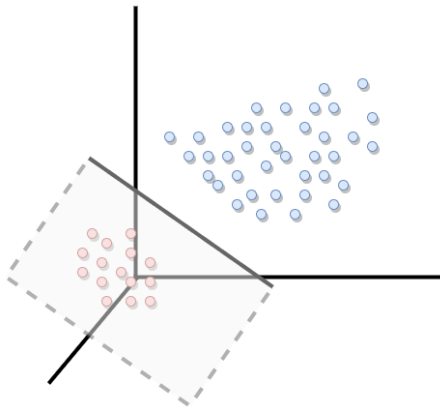
$$\Phi : X \rightarrow \mathfrak{F}$$

onde, X é o espaço de entrada, \mathfrak{F} o espaço de característica e Φ *kernel* usado

- A escolha apropriada de Φ permite que o conjunto de treinamento mapeado em \mathfrak{F} possa ser separado por uma SVN linear

SVM não-linear

- Existem duas condições que devem ser satisfeitas:
 - 1 A transformação não seja linear
 - 2 O quantidade de dimensões no espaço de característica seja suficientemente alta



SVM não-linear

Importante

Ao utilizar SVM não-linear, os novos exemplos que serão apresentados para a classificação pela SVM também são mapeados para o espaço de característica.

Esse mapeamento deve ser realizado pelo mesmo *kernel* antes da classificação.

Sumário

- 1 Recapitulação
- 2 Introdução
- 3 Vetor de Suporte
- 4 SVM linear
 - SVM com margens rígidas
 - SVM com margens suaves
- 5 SVM não-linear
- 6 SVM para problemas de regressão

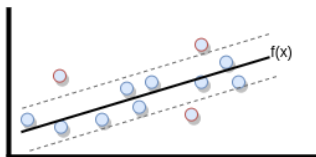
SVM para regressão

- É possível realizar adaptações para abordar problemas de regressão
- Nesse caso, o objetivo é estimar uma função $f(x)$ que produza como saída valores contínuos para os dados de entrada recebidos
- Os valores devem ter desvios máximos de ϵ , de acordo com os valores desejados
- Além disso, os valores de saída devem apresentar valores com variações uniformes e regulares

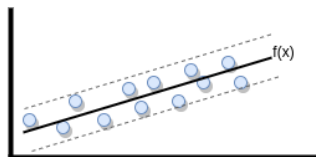
SVM para regressão

- Exemplo de uso para uma função linear:

Margens suaves



Margens rígidas



SVM para regressão não-linear

- Da mesma forma que para problemas de classificação não-lineares, a regressão não-linear pode ser realizada utilizando *kernels*
- Assim, um problema não-linear é mapeada para um espaço linear com mais dimensões e a função $f(x)$ é estimada
- Note que, os parâmetros de entrada são mapeados para o espaço de característica, mas os valores de resposta permanecem os mesmos

SCC 0270 – Redes Neurais e Aprendizado Profundo

Máquinas de Vetores de Suporte

Support Vector Machines – SVM

Prof. Bruno S. Faíçal

e-mail: bsfaical@alumni.usp.br

PAE: José Pedro Ribeiro Belo

e-mail: josepedroribeirobelo@gmail.com

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo