

Avaliação de métodos não paramétricos para predição em modelos aditivos

Orientador(^a): Prof^o Dr^o George Lucas Moraes Pezzott

Aluno(^a) : Marco Aurelio Valles Leal

RA: 103159

Maringá, 10 de maio de 2022

Conteúdo

Introdução

Objetivos

Metodologia

Resultados e Discussão

Aplicações

Introdução

Introdução

- ▶ A **análise de regressão** tem como objetivo descrever a relação entre uma variável resposta de interesse (y) e um conjunto de p variáveis preditoras. A variável resposta está relacionada às covariáveis na seguinte **relação linear**:

$$y = \beta_0 + \sum_{j=1}^p \beta_j x_j + \varepsilon \quad (1)$$

- ▶ A relação existente entre a variável resposta e cada uma das covariáveis **pode não ser linear**;
- ▶ Os **modelos aditivos** são um caso particular de uma classe mais geral denominada modelos aditivos generalizados (HASTIE & TIBSHIRANI, 1990), definido por,

$$y = \alpha + \sum_{j=1}^p f_j(x_j) + \varepsilon \quad (2)$$

- ▶ Observa-se que um modelo de regressão linear múltiplo é obtido adotando $f_j(x_j) = \beta_j x_j$;
- ▶ As funções f_j devem ser estimadas por meio de **suavizadores**, que estimam uma tendência menos variável e descreve sua dependência em relação à variável resposta.

Objetivos

Objetivos

- ▶ O objetivo deste trabalho é estudar algumas das principais técnicas de estimação do modelo aditivo no contexto não paramétrico.
- ▶ Em particular, serão abordados os suavizadores *Kernel*, *Loess* e *Splines* de regressão considerando apenas uma covariável, e seus desempenhos serão comparados com foco principal na predição.
 - ▶ Apresentar algumas técnicas de suavização para estimar funções não paramétricas presentes nos modelos aditivos;
 - ▶ Introduzir uma métrica para estimar o parâmetro de suavização de cada técnica;
 - ▶ Adotar uma métrica para avaliar a qualidade de estimação e predição;
 - ▶ Realizar um estudo de simulação para comparar a qualidade do ajuste e predição dos modelos em alguns cenários, tendo em vista os modelos aditivos;
 - ▶ Aplicar a metodologia estudada a um conjunto de dados reais, comparando modelos e técnicas de estimação e predição.

Metodologia

Metodologia - Suavizadores

- ▶ Considerar-se-á para o estudo da relação entre uma variável resposta (y) e uma variável explicativa (x) a partir de um modelo aditivo (HASTIE & TIBSHIRANI, 1990) da seguinte forma,

$$y = \alpha + f(x) + \varepsilon \quad (3)$$

- ▶ A função $f(x)$ do componente sistemático pode ser estimada a partir de um suavizador (*smoother*).
- ▶ No caso de uma variável preditora é chamado de suavizador em diagrama de dispersão.
- ▶ Os suavizadores possuem dois usos principais:
 - ▶ **Descrição;**
 - ▶ **Estimação.**

Metodologia - Suavizadores

- ▶ O Suavizador mais simples é o caso dos preditores **categóricos**;
- ▶ Para suavizar y podemos simplesmente **realizar as médias** de seus valores cada categoria.
- ▶ Este processo captura a **tendência** de y em x .
- ▶ Em relação as médias realizadas nas vizinhanças em torno do valor alvo, têm-se duas decisões a serem tomadas:
- ▶ O quão grande a vizinhança deve ser;
- ▶ Como realizar a média dos valores da resposta y em cada vizinhança.
- ▶ O tamanho da **vizinhança** a ser tomada é, normalmente, expressa em forma de um parâmetro (**parâmetro suavizador**).
- ▶ A questão de como realizar a média em uma vizinhança é a questão de qual **tipo de suavizador a utilizar**;

Metodologia - Suavizadores

- ▶ O suavizador **bin**;
- ▶ A média móvel (*running mean*);
- ▶ Linha Móvel;
- ▶ Suavizador Loess, também chamado de *Lowess*, essa técnica pode ser vista como uma linha móvel com pesos locais;
- ▶ Os pesos w_i são designados para cada ponto contidas nos vizinhos mais próximos, usando a função de peso tri-cúbica:

$$w_i = W\left(\frac{|x_0 - x_i|}{\Delta(x_0)}\right)$$

onde

$$W(u) = \begin{cases} (1 - u^3)^3, & 0 \leq u \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

- ▶ Considerando a i -ésima observação, ajusta-se uma regressão linear, levando em conta a função de peso tri-cúbica, que deve ser incorporada no modelo. A estimativa para $s(x_i)$, para i -ésima observação, será o valor predito de x_i em relação a seu i -ésimo modelo de regressão linear.

Metodologia - Suavizadores

- ▶ Um suavizador *kernel* usa **pesos que decrescem suavemente** enquanto se distância do ponto de interesse x_0 .
- ▶ Um suavizador *Kernel* pode ser definido da forma

$$\hat{y}_i = \frac{\sum_{j=1}^n y_j K\left(\frac{x_i - x_j}{b}\right)}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{b}\right)} \quad (4)$$

- ▶ Um *Spline* pode ser visto como uma função definida por um polinômio por partes.
- ▶ Pontos distintos são escolhidos no intervalo das observações (nós) e um **polinômio** é definido para **cada intervalo**.
- ▶ Além disso, o *spline* permite modelar um comportamento atípico dos dados, o que não seria possível com apenas uma função.
- ▶ Existem várias diferentes configurações para um *spline*, mas uma escolha popular é o ***spline* cúbico**;
- ▶ Quando trabalha-se com *splines*, existe uma dificuldade em escolher a **localização e quantidade ideal dos nós**.

Metodologia - Seleção de parâmetros - Função de Risco

- ▶ Para elaborar boas funções de predição, cria-se um critério para mensurar o desempenho que determinada função predição; valendo-se, por exemplo, do do risco quadrático (IZBICKI E SANTOS, 2020);

$$R_{pred}(g) = E [(Y - g(X))^2] . \quad (5)$$

- ▶ Constata-se que (X, Y) é uma observação nova não utilizada ao se estimar g . Sendo assim, melhor será a função de predição g , quanto **menor for o risco**;
- ▶ Usualmente ajusta-se distintos modelos para a função de regressão e encontrar qual deles apresenta um **melhor poder preditivo**;
- ▶ O método de seleção de modelos pretende selecionar uma **boa função** g .
- ▶ O risco observado, conhecido como erro quadrático médio em **relação aos dados**, e determinado por,

$$EQM(g) = \frac{1}{n} E \sum_{i=1}^n [(Y_i - g(X_i))^2] , \quad (6)$$

Metodologia - Seleção de parâmetros - Validação Cruzada

- ▶ Usualmente, é comum dividir os dados em dois conjuntos, um de treinamento e validação;
 - ▶ Algumas variações podem ser realizadas como o processo *k-fold cross validation*, ou ainda, *leave-one-out cross validation* (LOOCV);
 - ▶ O processo para seleção do melhor parâmetro de suavização por meio do *leave one out cross-validation* é descrito:
1. Supondo o parâmetro suavizador, denotado por p (tamanho do span ou número de nós), para cada valor possível de seu domínio faça:
 - 1.1 Considerando um conjunto de dados de tamanho n , para cada observação contidas no conjunto de dados faça:
 - 1.1.1 Divida o conjunto de dados, em dados de treino e validação;
 - 1.1.2 Construa (ajuste) o modelo utilizando apenas os dados de treino;
 - 1.1.3 Utilize o modelo para prever o valor da resposta ($\hat{y}_i = g(x_i)$), considerando a observação que compoe os dados de validação e calcule a distância $(y_i - \hat{y}_i)^2$.
 - 1.2 Repita o processo (1.1) n vezes até que todas as observações sejam "vistas" como dados de validação.

$$EQM(p) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

2. Repita a etapa (1), para todo o domínio do parâmetro suavizador.

Metodologia - Seleção das técnicas de suavização

- ▶ Após selecionado o melhor parâmetro de suavização, ter-se-á um parâmetro considerado ótimo;
- (i) EQM_c : Erro quadrático médio "completo": Após a escolha do parâmetro de suavização e tendo em vista todos as observações do conjuntos dados, ajusta-se o modelo. Para cada observação, considere a predição $\hat{y}_i = g(x_i)$, para todo $i = 1, 2, \dots, n$. Em seguida, calcula-se o EQM:

$$EQM_c = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

- (ii) EQM_{loocv} : Erro quadrático médio "LOOCV": Essa métrica será exatamente o EQM durante a escolha do parâmetro suavizador.

Resultados e Discussão

Estudo de Simulação

- ▶ Serão utilizadas simulações de dados;
- ▶ Analisar-se-á as performances dos resultados obtidos, de quatro técnicas de suavização sendo elas: o suavizador de *kernel*, *Loess*, *splines* de regressão de linear e cúbico.;
- ▶ o parâmetro suavizador é indicado, usualmente, pela largura da banda ou *span*.
- ▶ Realizar-se-ão ajustes para o primeiro cenário, considerando distintos parâmetros de suavização;
- ▶ Em seguida, adotando o método de *data splitting*, *leave one out cross-validation*;
- ▶ Posteriormente, este procedimento será repetido para cada cenário em mil amostras;
- ▶ Será realizada escolha do melhor parâmetro de suavização selecionados por meio da métrica EQM_{loocv} .
- ▶ Em seguida, comparar-se-á dentro dos mesmos cenários, avaliando por meio da métrica EQM_c ;
- ▶ Vale ressaltar, que serão empregados dois comportamentos, um proveniente de uma função senoidal e outra de uma função Gamma: Cenário 1 e Cenário 2.

Cenário 1 - Introdução

- ▶ Valores de X uma sequência de 0 a 50;



$$y = 10 + 5\sin\pi\frac{x}{24} + \varepsilon \quad (7)$$

- ▶ ε é um termo aleatório, normalmente, distribuído com média zero e variância constante;
- ▶ Tamanhos amostrais utilizados serão iguais a 150, 250 e 350;
- ▶ Valores de desvio padrão, utilizado nos erros, 0.5, 1 e 2.

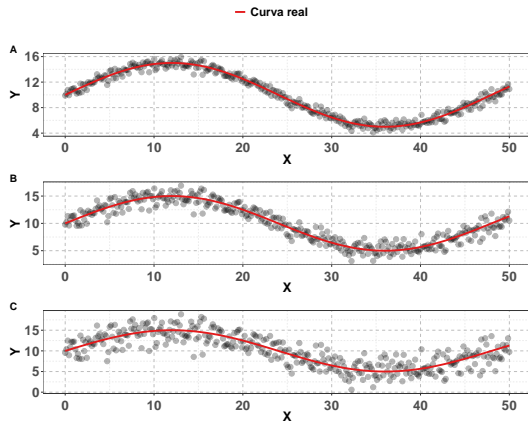


Figura 1: Gráfico de dispersão dos dados simulados e curva real, para o Cenário 1. (A) DP = 0,5, (B) DP = 1 e (C) DP = 2.

Cenário 1 - Ajustes suavizados

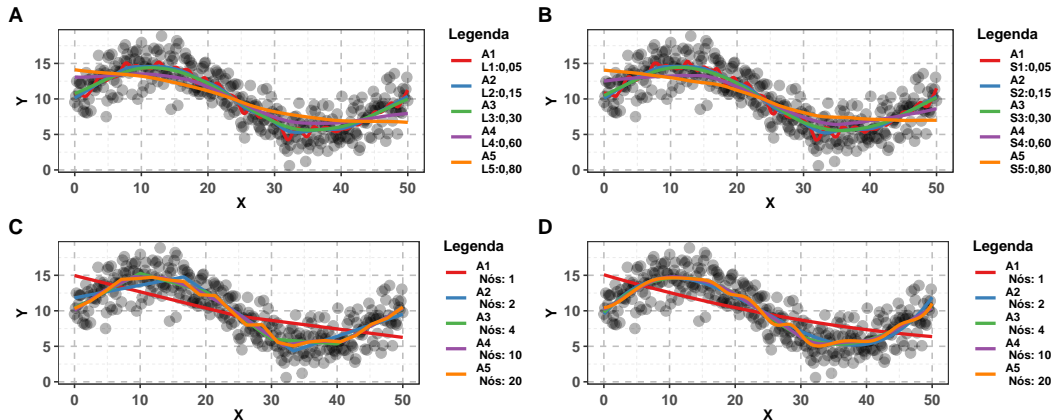


Figura 2: Comparação entre diferentes ajustes (Cenário 1), com parâmetros de suavização distintos, considerando os suavizadores, (A) Kernel, (B) Loess, (C) Splines de Regressão Linear e (D) Splines de Regressão Cúbico.

Cenário 1 - Seleção dos parâmetros EQM_{loocv}

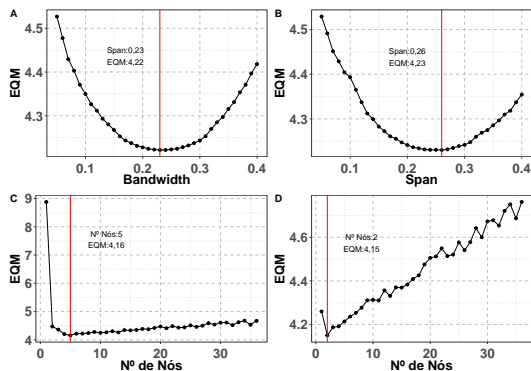


Figura 3: Erro quadrático médio (EQM_{loocv}) versus parâmetro de suavização (Cenário 1) pós aplicação do Leave One Out Cross-Validation. (A) Kernel, (B) Loess, (C) Splines de Regressão Linear e (D) Splines de Regressão Cúbico.

Tabela 1: Erro Quadrático Médio (EQM_{loocv}), para os suavizadores Kernel, Loess e Splines de Regressão Linear e Cúbico.

Suavizador	Parâm. Suavizador	EQM
Kernel	0,23	4,22
Loess	0,26	4,23
Sp. Reg. Linear	5	4,16
Sp. Reg. Cúbico	2	4,15

Tabela 2: Erro Quadrático Médio (EQM_c), para os suavizadores Kernel, Loess e Splines de Regressão Linear e Cúbico.

Suavizador	Parâm. Suavizador	EQM
Kernel	0,23	0,2652
Loess	0,26	0,2634
Sp. Reg. Linear	5	0,2674
Sp. Reg. Cúbico	2	0,2723

Cenário 1 - Resultados do EQM_{loocv} em 1000 amostras

Tabela 3: Percentual do Erro quadrático médio EQM_{loocv} mínimo, obtidos por meio de aplicação do Procedimento 1, para seleção do melhor parâmetro de suavização, considerando cada suavizador em 1000 amostras.

Sub-Cenário	Tamanho	Desvio Padrão	Kernel	Loess	Sp. Reg. Linear	Sp. Reg. Cúbico
1	150	0,5	0,9%	5,3%	19,7%	74,1%
2	150	1,0	1,4%	8,4%	31,1%	59,1%
3	150	2,0	1,4%	10,8%	48,0%	39,8%
4	250	0,5	0,7%	3,8%	16,9%	78,6%
5	250	1,0	1,1%	7,0%	23,5%	68,4%
6	250	2,0	1,7%	13,1%	35,8%	49,4%
7	350	0,5	0,8%	3,7%	15,8%	79,7%
8	350	1,0	1,3%	7,0%	20,5%	71,2%
9	350	2,0	1,8%	11,7%	33,5%	53,0%

Cenário 1 - Comportamento dos EQM_{loocv}

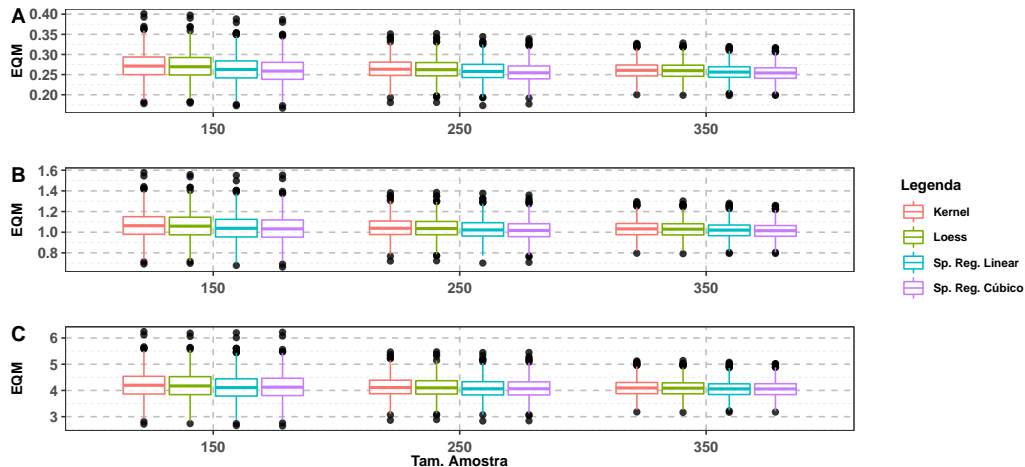


Figura 4: Comparação do erro quadrático para as 1000 amostras para cada suavizador. (A) $DP = 0.5$, (B) $DP = 1$ e (C) $DP = 2$

Resultados do $EQMc$ em 1000 amostras, pós seleção dos parâmetros

Tabela 4: Percentual do Erro quadrático médio EQM_{loocv} mínimo, considerando os ajustes provenientes dos melhores parâmetros obtidos por meio de aplicação do Procedimento 1, para cada suavizador em 1000 amostras.

Sub-Cenário	Tamanho	Desvio Padrão	Kernel	Loess	Sp. Reg. Linear	Sp. Reg. Cúbico
1	150	0,5	65,3%	15,4%	11,7%	7,6%
2	150	1,0	61,3%	11,8%	14,6%	12,3%
3	150	2,0	63,2%	11,2%	11,3%	14,3%
4	250	0,5	67,2%	17,1%	11,0%	4,7%
5	250	1,0	66,1%	14,0%	12,7%	7,2%
6	250	2,0	56,2%	12,5%	13,9%	17,4%
7	350	0,5	79,9%	8,8%	8,7%	2,6%
8	350	1,0	63,3%	16,8%	13,1%	6,8%
9	350	2,0	56,2%	12,4%	13,8%	17,6%

Comportamento para os EQM_c

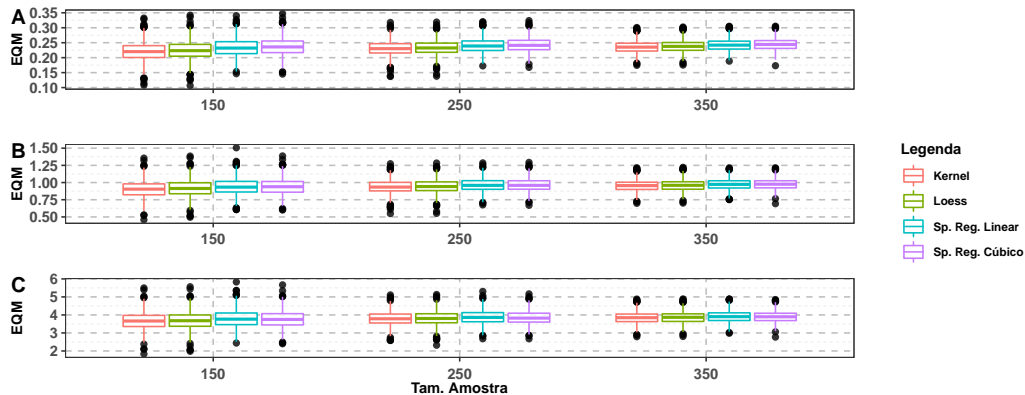


Figura 5: Comparação do erro quadrático para as 1000 amostras para cada suavizador. (A) $DP = 0.5$, (B) $DP = 1$ e (C) $DP = 2$

Cenário 2 - Introdução

- ▶ Valores de X uma sequência de 0.1 a 2;
- ▶ $y = f(x) + \varepsilon$, com $f(x) \sim \text{Gamma}(6, 10)$;
- ▶ $\varepsilon \sim N(0, \sigma^2)$;
- ▶ Tamanhos amostrais utilizados serão iguais a 150, 250 e 350;
- ▶ Valores de desvio padrão, utilizado nos erros, 0.05, 0.1 e 0.15.

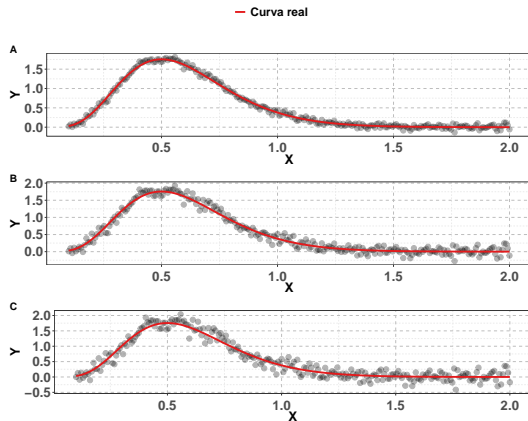


Figura 6: Gráfico de dispersão dos dados simulados e curva real, para o Cenário 2. (A) DP = 0,05, (B) DP = 0,10 e (C) DP = 0,15.

Cenário 2 - Resultados dos EQM_{loocv}

Tabela 5: Percentual do Erro quadrático médio EQM_{loocv} mínimo, considerando os ajustes provenientes dos melhores parâmetros obtidos por meio de aplicação do Procedimento 1, para cada suavizador em 1000 amostras.

Sub-Cenário	Tamanho	Desvio Padrão	Kernel	Loess	Sp. Reg. Linear	Sp. Reg. Cúbico
1	150	0,05	0,6%	2,6%	7,4%	89,4%
2	150	0,10	0,7%	6,3%	13,5%	79,5%
3	150	0,15	0,7%	9,1%	20,6%	69,6%
4	250	0,05	0,4%	1,7%	8,2%	89,7%
5	250	0,10	0,7%	3,4%	10,3%	85,6%
6	250	0,15	0,8%	6,5%	12,7%	80,0%
7	350	0,05	0,2%	2,0%	5,9%	91,9%
8	350	0,10	0,6%	2,5%	8,9%	88,0%
9	350	0,15	1,0%	4,7%	11,4%	82,9%

Cenário 2 - Comportamento para os EQM_{loocv}

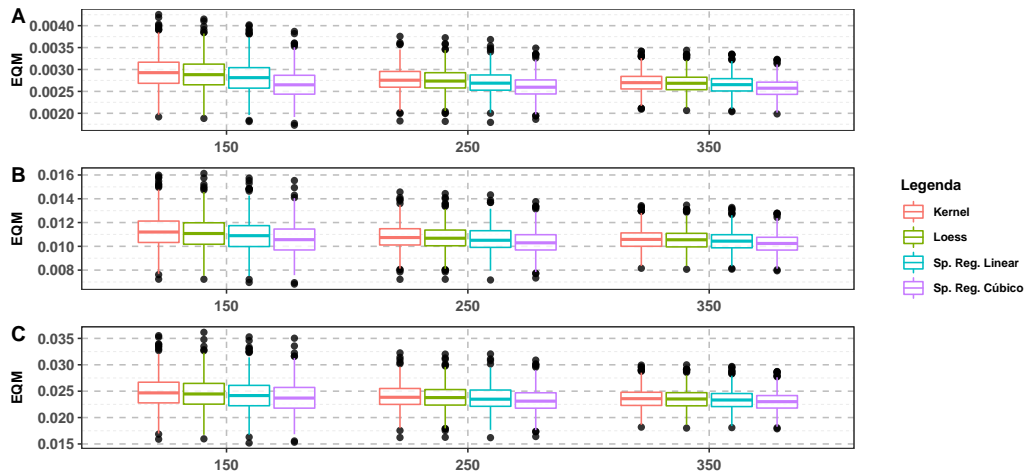


Figura 7: Comparação do erro quadrático médio (EQM_{loocv}) em 1000 amostras para cada suavizador. (A) $DP = 0.05$, (B) $DP = 0.1$ e (C) $DP = 0.15$

Cenário 2 - Resultados para o EQM_c , pós seleção dos parâmetros

Tabela 6: Percentual do Erro quadrático médio EQM_{loocv} mínimo, considerando os ajustes provenientes dos melhores parâmetros obtidos por meio de aplicação do Procedimento 1, para cada suavizador em 1000 amostras.

Sub-Cenário	Tamanho	Desvio Padrão	Kernel	Loess	Sp. Reg. Linear	Sp. Reg. Cúbico
1	150	0,5	90,1%	6,0%	3,8%	0,1%
2	150	1,0	71,1%	9,3%	16,3%	3,3%
3	150	2,0	49,1%	13,5%	25,6%	11,8%
4	250	0,5	91,9%	4,2%	3,5%	0,4%
5	250	1,0	79,9%	7,6%	10,7%	1,8%
6	250	2,0	61,9%	10,6%	20,3%	7,2%
7	350	0,5	91,9%	3,7%	3,7%	0,7%
8	350	1,0	82,0%	7,6%	8,7%	1,7%
9	350	2,0	65,4%	10,3%	20,3%	4,0%

Cenário 2 - Comportamento para os EQM_c

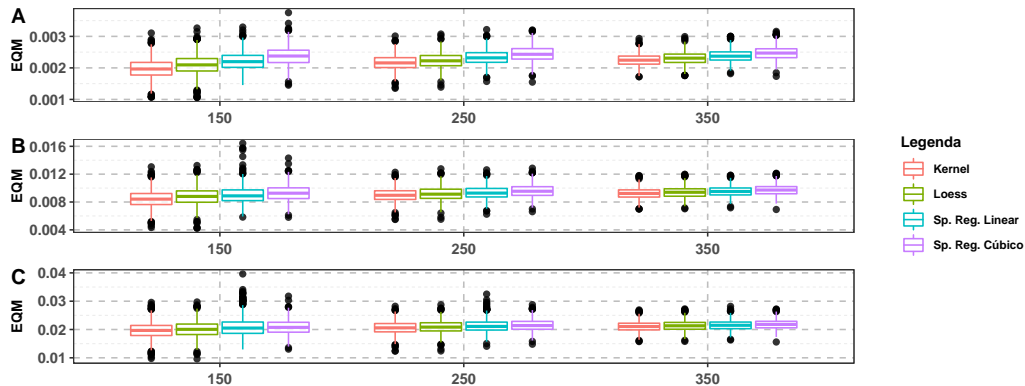


Figura 8: Comparação do erro quadrático médio (EQM_c), para as 1000 amostras para cada suavizador. (A) $DP = 0.05$, (B) $DP = 0.1$ e (C) $DP = 0.15$

Aplicações

Aplicação 1

- ▶ Os dados é referente à expansão térmica de cobre.
- ▶ A variável resposta é o coeficiente de expansão térmica e a variável preditora é a temperatura em graus kelvin.
- ▶ Considerando o modelo aditivo da seguinte forma

$$y = \alpha + f(x) + \varepsilon,$$

- ▶ onde os erros ε são independentes, com $E(\varepsilon) = 0$ e $var(\varepsilon) = \sigma^2$.

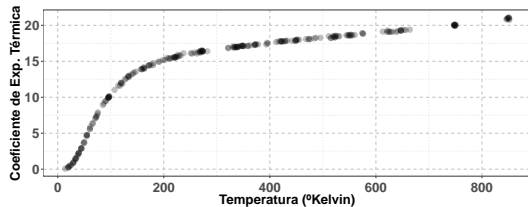


Figura 9: Gráfico de dispersão referente ao estudo de expansão térmica do cobre versus a temperatura em graus Kelvin.

Avaliando os EQM_{loocv} e os EQM_c

Tabela 7: Erro Quadrático Médio (EQM_{loocv}) para os suavizadores Loess, Kernel e Splines de Regressão Linear e Cúbico (Aplicação 1).

Suavizador	Parâm. Suavizador	EQM
Kernel	0,1	0,0091
Loess	0,1	0,0078
Sp. Reg. Linear	20,0	0,0078
Sp. Reg. Cúbico	14,0	0,0065

Tabela 8: Erro Quadrático Médio (EQM_c) para os suavizadores Loess, Kernel e Splines de Regressão Linear e Cúbico.

Suavizador	Parâm. Suavizador	EQM
Kernel	0.1	0,000200
Loess	0.1	0,000211
Sp. Reg. Linear	20	0,000211
Sp. Reg. Cúbico	14	0,000171
Polinômio Cúbico		0,031052

Comparação dos Ajustes

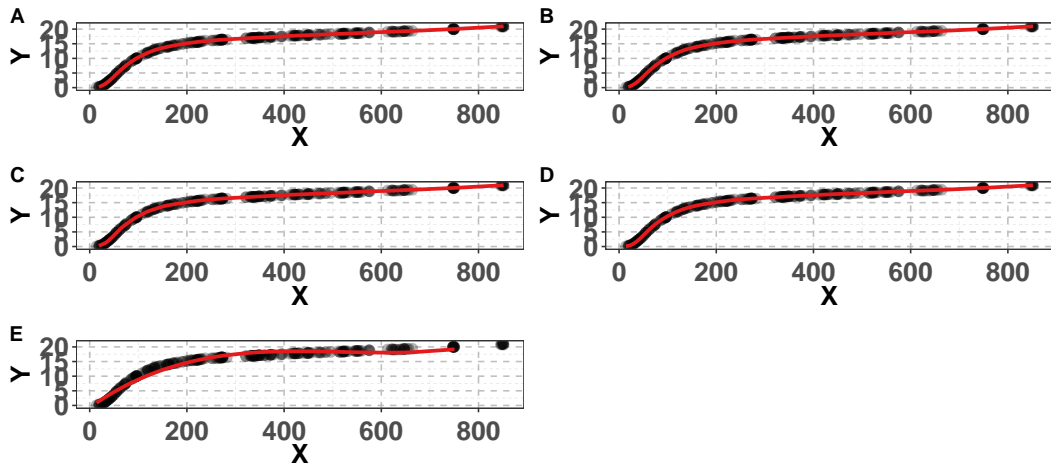


Figura 10: Comparação entre os ajustes, considerando parâmetros de suavização obtidos por meio da validação cruzada. (A) Kernel - Parâm. 0,1. (B) Loess - Parâm. 0,1. (C) Splines de Regressão Linear - Parâm. 20. (D) Splines de Regressão Cúbico - Parâm. 14. (E) Regressão Polinômial Cúbica.

Aplicação 2

- ▶ Os dados são referentes ao preço mediano de casas de Boston.
- ▶ A variável resposta y , $medv$, que representa o valor mediano das casas ocupadas pelos proprietários em 1000 dolars.
- ▶ Para a variável preditora x , $lstat$, representa o percentual de status baixo da população.
- ▶ Considerando o modelo aditivo da seguinte forma,

$$y = \alpha + f(x) + \varepsilon,$$

- ▶ onde os erros ε são independentes, com $E(\varepsilon) = 0$ e $var(\varepsilon) = \sigma^2$. A $f(x)$.

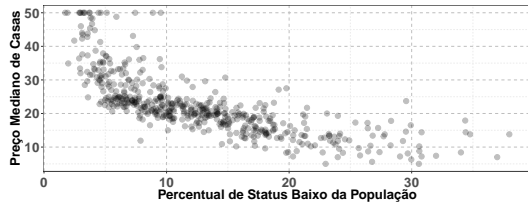


Figura 11: Gráfico de dispersão referente ao preço mediano de casas versus o percentual de status baixos na população.

Avaliando os EQM_{loocv} e os EQM_c

Tabela 9: Erro Quadrático Médio (EQM_{loocv}) para os suavizadores Loess, Kernel e Splines de Regressão Linear e Cúbico (Aplicação 1).

Suavizador	Parâm. Suavizador	EQM
Kernel	0,21	27,1909
Loess	0,18	27,2788
Sp. Reg. Linear	5,00	27,1191
Sp. Reg. Cúbico	5,00	27,3552

Tabela 10: Erro Quadrático Médio (EQM_c) para os suavizadores Loess, Kernel e Splines de Regressão Linear e Cúbico.

Suavizador	Parâm. Suavizador	EQM
Kernel	0.21	0,307130
Loess	0.18	0,308202
Sp. Reg. Linear	5	0,315720
Sp. Reg. Cúbico	5	0,314639
Polinômio Cúbico		0,342152

Comparação dos ajustes

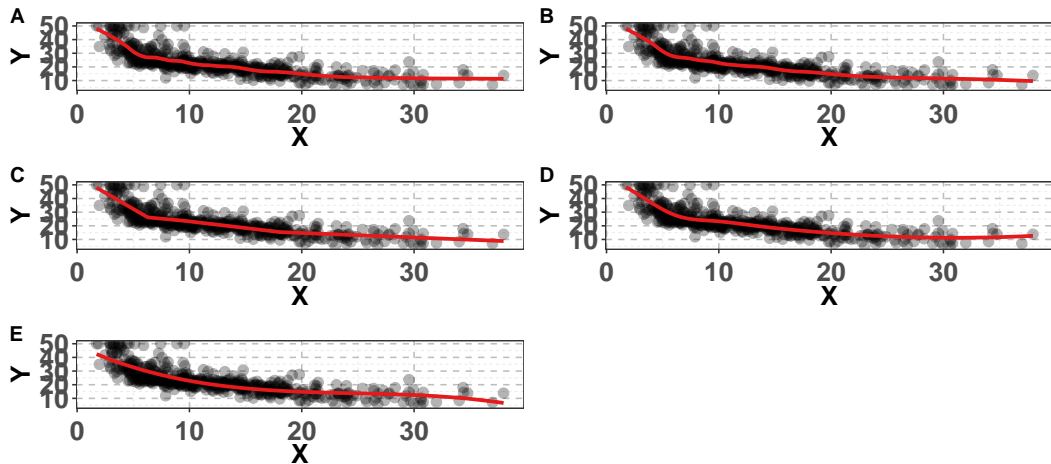


Figura 12: Comparação entre os ajustes, considerando parâmetros de suavização obtidos por meio da validação cruzada. (A) Kernel - Parâm. 0,21. (B) Loess - Parâm. 0,18. (C) Splines de Regressão Linear - Parâm. 5 (D) Splines de Regressão Cúbico - Parâm. 5 (E) Regressão Polinômial Cúbica.

Considerações Finais

- ▶ As técnicas de suavização *kernel*, *loess* e splines de regressão, em particular, os de grau um e grau três foram apresentadas, que são utilizadas para estimar as funções presentes em modelos aditivos.
- ▶ Foi introduzido um método para obtenção do melhor parâmetro suavizador, realizado por meio de método de validação cruzada.
- ▶ Duas métricas foram abordadas: uma para verificar a qualidade de predição, erro quadrático médio obtidos por validação cruzada, *leave one out cross validation* (EQM_{loocv}) e a outra para apurar a qualidade dos ajustes, erro quadrático médio completo (EQM_c).
- ▶ Para validar a metodologia estudada, foram realizadas análises em dados simulados e dados reais.
- ▶ Ainda, verificou-se que o ajuste mais adequado para descrever o comportamento dos dados não obtém necessariamente o melhor poder preditivo.

Considerações Finais

- ▶ Por meio das métricas, (EQM_{loocv}) e (EQM_c) , o suavizador que obteve o melhor poder preditivo foi o *splines* de regressão cúbico. Levando em consideração o método com melhor qualidade de ajustes, os suavizadores com *kernel* se destacou em ambos os cenários.
- ▶ Ressalta-se que para a Aplicação 1, a técnica que obteve o melhor poder preditivo e o melhor ajuste fora o *splines* de regressão cúbico.
- ▶ Para a Aplicação 2, o suavizador que se destacou por obter o melhor poder preditivo (menor EQM_{loocv} entre os suavizadores) foi *splines* de regressão linear. Em contrapartida, o que denotou melhor qualidade de ajuste (menor EQM_c dentre os suavizadores) foi o método *kernel* gaussiano.

Considerações Finais - Sugestões para trabalhos futuros

- ▶ Métricas para validação da qualidade de predição e adequabilidade dos modelos podem ser adotadas.
- ▶ Técnicas que podem ser adotadas para seleção dos parâmetros de suavização.
- ▶ Para os suavizadores *splines* de regressão, além da seleção da quantidade de nós, a localização dos nós pode ser avaliada a fim de obter um melhor ajuste.
- ▶ Ainda as discussões podem ser extendidas, quando mais de uma covariável está disponível para prever a resposta. Frequentemente, utiliza-se o algoritmo de retroajuste (*backfitting*, HASTIE & TIBSHIRANI, 1990) para estimar cada função suave f_j em um cenário não paramétrico.

Referências

BUJA, A., HASTIE, T. & TIBSHIRANI, R. (1989). **Linear smoothers and additive models.** The Annals of Statistics, 17, 453-510.

CLEVELAND, W. S. (1979). **Robust locally weighted regression and smoothing scatterplots.** Journal of the American Statistical Association, 74, 829-836.

DELICADO, P., 2008 **Curso de Modelos no Paramétricos** p. 200.

EUBANK, R. L(1999) **Nonparametric Regression and Spline Smoothing.** Marcel Dekker, 2o edição. Citado na pág. 1, 2, 29

FAHRMEIR, L. & TUTZ, G. (2001) **Multivariate Statistical Modelling Based on Generalized Linear Models.** Springer, 2o edição. Citado na pág. 15

GREEN, P. J. & YANDELL, B. S. (1985) **Semi-parametric generalized linear models.** Lecture Notes in Statistics, 32:4455. Citado na pág. 15

Referências

GREEN P. J. & SILVERMAN B. W. (1994). **Nonparametric regression and generalized linear models: a roughness penalty approach**. Chapman & Hall, London.

HASTIE, T. J. & TIBSHIRANI, R. J. (1990). **Generalized additive models**, volume 43. Chapman and Hall, Ltd., London. ISBN 0-412-34390-8.

MONTGOMERY, D. C. & PECK, E. A. & VINING, G. G. **Introduction to Linear Regression Analysis**. 5th Edition. John Wiley & Sons, 2012.

IZBICK, R. & SANTOS, T. M. **Aprendizado de máquina: uma abordagem estatística**. ISBN 978-65-00-02410-4.

TEAM, R. CORE. R: **A language and environment for statistical computing**. (2013).

Fim

Obrigado!

”Sem números, não há vantagem nem probabilidades; sem vantagens e probabilidades, o único meio de lidar com o Risco é apelar para os deuses e o destino. Sem números, o RISCO é uma questão de pura CORAGEM.” (Peter L. Bernstein).