

CE092- Extensões do modelo de regressão

Modelos aditivos generalizados

Cesar Augusto Taconeli

14 de setembro, 2026

O que vamos estudar nesta aula

O que vamos estudar nesta aula

- A extensão do modelo linear generalizado ao modelo aditivo generalizado;

O que vamos estudar nesta aula

- A extensão do modelo linear generalizado ao modelo aditivo generalizado;
- Funções suaves, penalização e EDF;

O que vamos estudar nesta aula

- A extensão do modelo linear generalizado ao modelo aditivo generalizado;
- Funções suaves, penalização e EDF;
- Ajuste, interpretação e diagnóstico dos GAMs;

O que vamos estudar nesta aula

- A extensão do modelo linear generalizado ao modelo aditivo generalizado;
- Funções suaves, penalização e EDF;
- Ajuste, interpretação e diagnóstico dos GAMs;
- Uso de múltiplos suavizadores, respostas não Gaussianas e efeitos bidimensionais.

Introdução

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:
 - **Splines de regressão cúbicos:** opção clássica para suavização univariada;

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:
 - **Splines de regressão cúbicos:** opção clássica para suavização univariada;
 - **Thin plate regression splines:** opção bastante geral e padrão no pacote mgcv;

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:
 - **Splines de regressão cúbicos:** opção clássica para suavização univariada;
 - **Thin plate regression splines:** opção bastante geral e padrão no pacote mgcv;
 - **P-splines:** splines combinados com penalização para controlar a suavidade;

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:
 - **Splines de regressão cúbicos:** opção clássica para suavização univariada;
 - **Thin plate regression splines:** opção bastante geral e padrão no pacote mgcv;
 - **P-splines:** splines combinados com penalização para controlar a suavidade;
 - **Suavizadores bidimensionais:** utilizados para representar efeitos conjuntos ou interações entre covariáveis, dentre outros.

- Modelos aditivos generalizados (GAMs) estendem os modelos lineares generalizados (GLMs), permitindo que relações não lineares entre as covariáveis e a média da resposta sejam representadas por funções suaves.
- Dentre as principais bases e estruturas de suavização disponíveis para modelos GAM, destacamos:
 - **Splines de regressão cúbicos:** opção clássica para suavização univariada;
 - **Thin plate regression splines:** opção bastante geral e padrão no pacote mgcv;
 - **P-splines:** splines combinados com penalização para controlar a suavidade;
 - **Suavizadores bidimensionais:** utilizados para representar efeitos conjuntos ou interações entre covariáveis, dentre outros.
- Independentemente da base utilizada, a complexidade das funções suaves é controlada por meio de **penalização**.

Exemplo 1 - Emissão de gases na combustão de etanol

Exemplo 1 - Emissão de gases na combustão de etanol

- Para fins de ilustração, vamos usar os dados da base `ethanol`, disponíveis na biblioteca `lattice` do R.

Exemplo 1 - Emissão de gases na combustão de etanol

- Para fins de ilustração, vamos usar os dados da base `ethanol`, disponíveis na biblioteca `lattice` do R.
- A base é composta por $n = 88$ registros e as variáveis que vamos analisar são as seguintes:

Exemplo 1 - Emissão de gases na combustão de etanol

- Para fins de ilustração, vamos usar os dados da base `ethanol`, disponíveis na biblioteca `lattice` do R.
- A base é composta por $n = 88$ registros e as variáveis que vamos analisar são as seguintes:
 - **E**: Razão de equivalência, uma medida da riqueza da mistura de ar e etanol combustível;

Exemplo 1 - Emissão de gases na combustão de etanol

- Para fins de ilustração, vamos usar os dados da base `ethanol`, disponíveis na biblioteca `lattice` do R.
- A base é composta por $n = 88$ registros e as variáveis que vamos analisar são as seguintes:
 - **E**: Razão de equivalência, uma medida da riqueza da mistura de ar e etanol combustível;
 - **NOx**: Concentração de óxidos de nitrogênio (NO e NO₂) emitidos, em microgramas/J.

Exemplo 1 - Emissão de gases na combustão de etanol

- Para fins de ilustração, vamos usar os dados da base `ethanol`, disponíveis na biblioteca `lattice` do R.
- A base é composta por $n = 88$ registros e as variáveis que vamos analisar são as seguintes:
 - **E**: Razão de equivalência, uma medida da riqueza da mistura de ar e etanol combustível;
 - **NOx**: Concentração de óxidos de nitrogênio (NO e NO₂) emitidos, em microgramas/J.
- O objetivo é ajustar um modelo de regressão considerando NOx a resposta e E o preditor (variável explicativa).

Exemplo 1 - Emissão de gases na combustão de etanol

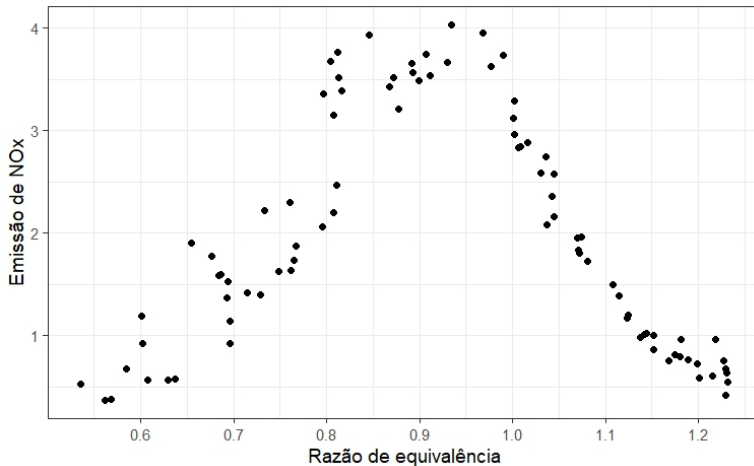


Figura 1: Relação entre concentração de óxidos de nitrogênio emitidos e razão de equivalência.

Exemplo 1 - Emissão de gases na combustão de etanol

- A relação entre a razão de equivalência e a emissão de NOx é claramente não linear.

Exemplo 1 - Emissão de gases na combustão de etanol

- A relação entre a razão de equivalência e a emissão de NOx é claramente não linear.
- A especificação de uma forma funcional paramétrica adequada para representar essa relação não é evidente.

Exemplo 1 - Emissão de gases na combustão de etanol

- A relação entre a razão de equivalência e a emissão de NO_x é claramente não linear.
- A especificação de uma forma funcional paramétrica adequada para representar essa relação não é evidente.
- Uma alternativa é representar essa relação por meio de uma função suave, cuja forma é estimada a partir dos próprios dados.

Exemplo 1 - Emissão de gases na combustão de etanol

- Ao invés de definir uma regressão paramétrica:

$$E(NO_x | E) = f(E; \beta),$$

como uma regressão polinomial ou não linear, propõe-se:

Exemplo 1 - Emissão de gases na combustão de etanol

- Ao invés de definir uma regressão paramétrica:

$$E(\text{NOx} | \mathbf{E}) = f(\mathbf{E}; \boldsymbol{\beta}),$$

como uma regressão polinomial ou não linear, propõe-se:

$$E(\text{NOx} | \mathbf{E}) = \beta_0 + s(\mathbf{E}),$$

em que $s(\cdot)$ é uma função suave, cuja forma é estimada a partir dos dados.

Modelos lineares generalizados (GLM)

Modelos lineares generalizados

- No contexto de análise de regressão, seja Y a variável resposta e $\mathbf{x}' = (x_1, x_2, \dots, x_k)$ o vetor de valores observados das k covariáveis.

Modelos lineares generalizados

- No contexto de análise de regressão, seja Y a variável resposta e $\mathbf{x}' = (x_1, x_2, \dots, x_k)$ o vetor de valores observados das k covariáveis.
- Um **modelo linear generalizado** é definido por:

$$Y|\mathbf{x} \sim f(\mu_{\mathbf{x}}, \phi)$$

$$g(\mu_{\mathbf{x}}) = \mathbf{x}'\boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k,$$

Modelos lineares generalizados

- No contexto de análise de regressão, seja Y a variável resposta e $\mathbf{x}' = (x_1, x_2, \dots, x_k)$ o vetor de valores observados das k covariáveis.
- Um **modelo linear generalizado** é definido por:

$$Y|\mathbf{x} \sim f(\mu_{\mathbf{x}}, \phi)$$

$$g(\mu_{\mathbf{x}}) = \mathbf{x}'\boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k,$$

em que $f(\cdot)$ representa alguma distribuição pertencente à **família exponencial**, $g(\cdot)$ a **função de ligação** e ϕ um parâmetro de dispersão.

- Dentre as principais distribuições pertencentes à família exponencial temos a binomial, Poisson, normal, gama, normal inversa e binomial negativa.

Modelos lineares generalizados

- Dentre as principais distribuições pertencentes à família exponencial temos a binomial, Poisson, normal, gama, normal inversa e binomial negativa.
- A título de ilustração, um primeiro caso particular importante dos GLMs é a **regressão linear**:

$$Y|\mathbf{x} \sim \text{Normal}(\mu_{\mathbf{x}}, \sigma^2)$$

$$\mu_{\mathbf{x}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k.$$

- Na modelagem de dados binários, a **regressão logística** é bastante usual:

$$Y|\mathbf{x} \sim \text{Bernoulli}(\pi_{\mathbf{x}})$$

$$\log\left(\frac{\pi_{\mathbf{x}}}{1 - \pi_{\mathbf{x}}}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k.$$

- O **modelo log-linear de Poisson** pode ser usado na modelagem de contagens:

$$Y|\mathbf{x} \sim \text{Poisson}(\mu_{\mathbf{x}})$$

$$\log(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k.$$

- O **modelo log-linear de Poisson** pode ser usado na modelagem de contagens:

$$Y|\mathbf{x} \sim \text{Poisson}(\mu_{\mathbf{x}})$$

$$\log(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k.$$

- Outros GLMs podem ser definidos alterando a distribuição da resposta e/ou a função de ligação.

Modelos aditivos generalizados (GAM)

Modelos aditivos generalizados (GAM)

- Num modelo aditivo generalizado, o efeito de uma ou mais covariáveis pode ser introduzido ao modelo por meio de funções suaves.

Modelos aditivos generalizados (GAM)

- Num modelo aditivo generalizado, o efeito de uma ou mais covariáveis pode ser introduzido ao modelo por meio de funções suaves.
- Na sequência é apresentada a formulação geral de um **modelo aditivo generalizado**:

$$Y|\mathbf{x} \sim f(\mu_{\mathbf{x}}, \phi)$$

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_q x_q + s_{q+1}(x_{q+1}) + \dots + s_k(x_k),$$

em que $s_{q+1}(\cdot), \dots, s_k(\cdot)$ são suavizadores aplicados a $x_{q+1}, \dots, x_k$, $f(\cdot)$, $g(\cdot)$ e ϕ são definidos como em um GLM.

Modelos aditivos generalizados (GAM)

- Num modelo aditivo generalizado, o efeito de uma ou mais covariáveis pode ser introduzido ao modelo por meio de funções suaves.
- Na sequência é apresentada a formulação geral de um **modelo aditivo generalizado**:

$$Y|\mathbf{x} \sim f(\mu_{\mathbf{x}}, \phi)$$

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_q x_q + s_{q+1}(x_{q+1}) + \dots + s_k(x_k),$$

em que $s_{q+1}(\cdot), \dots, s_k(\cdot)$ são suavizadores aplicados a $x_{q+1}, \dots, x_k$, $f(\cdot)$, $g(\cdot)$ e ϕ são definidos como em um GLM.

- Observe que podemos ter termos paramétricos e não paramétricos compondo o preditor linear.

- GLM:

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

- GLM:

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

- GAM:

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_q x_q + s_{q+1}(x_{q+1}) + \dots + s_k(x_k)$$

- GLM:

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

- GAM:

$$g(\mu_{\mathbf{x}}) = \beta_0 + \beta_1 x_1 + \dots + \beta_q x_q + s_{q+1}(x_{q+1}) + \dots + s_k(x_k)$$

- A distribuição da resposta e a função de ligação permanecem; flexibiliza-se a forma como as covariáveis entram no preditor.

Modelos aditivos generalizados

- Em um GAM, cada termo $s_j(x_j)$ representa o efeito suave da covariável x_j , mantendo-se fixos os demais termos do modelo.

Modelos aditivos generalizados

- Em um GAM, cada termo $s_j(x_j)$ representa o efeito suave da covariável x_j , mantendo-se fixos os demais termos do modelo.
- As funções suaves são representadas por combinações de funções de base:

$$s_j(x_j) = \sum_{r=1}^{K_j} \beta_{jr} b_{jr}(x_j)$$

Modelos aditivos generalizados

- Em um GAM, cada termo $s_j(x_j)$ representa o efeito suave da covariável x_j , mantendo-se fixos os demais termos do modelo.
- As funções suaves são representadas por combinações de funções de base:

$$s_j(x_j) = \sum_{r=1}^{K_j} \beta_{jr} b_{jr}(x_j)$$

- A complexidade das funções suaves é controlada por **penalização**.

- Considerando, para simplificar, um único termo suave, o ajuste de um GAM baseia-se na maximização da **verossimilhança penalizada**:

$$\ell_p(\boldsymbol{\beta}) = \ell(\boldsymbol{\beta}) - \frac{1}{2}\lambda\boldsymbol{\beta}'\mathbf{S}\boldsymbol{\beta}.$$

- Considerando, para simplificar, um único termo suave, o ajuste de um GAM baseia-se na maximização da **verossimilhança penalizada**:

$$\ell_p(\boldsymbol{\beta}) = \ell(\boldsymbol{\beta}) - \frac{1}{2}\lambda\boldsymbol{\beta}'\mathbf{S}\boldsymbol{\beta}.$$

- O primeiro termo recompensa o bom ajuste aos dados; o segundo penaliza funções excessivamente complexas.

Modelos aditivos generalizados

- Considerando, para simplificar, um único termo suave, o ajuste de um GAM baseia-se na maximização da **verossimilhança penalizada**:

$$\ell_p(\beta) = \ell(\beta) - \frac{1}{2}\lambda\beta'\mathbf{S}\beta.$$

- O primeiro termo recompensa o bom ajuste aos dados; o segundo penaliza funções excessivamente complexas.
- A penalização não impede que a curva seja complexa; ela exige que essa complexidade seja justificada por uma melhora suficiente no ajuste.

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.
- A definição do parâmetro de suavização é crucial para o ajuste do GAM:

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.
- A definição do parâmetro de suavização é crucial para o ajuste do GAM:
 - Valores pequenos de λ produzem maior flexibilidade (menor suavização);

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.
- A definição do parâmetro de suavização é crucial para o ajuste do GAM:
 - Valores pequenos de λ produzem maior flexibilidade (menor suavização);
 - Valores grandes de λ produzem menor flexibilidade (maior suavização);

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.
- A definição do parâmetro de suavização é crucial para o ajuste do GAM:
 - Valores pequenos de λ produzem maior flexibilidade (menor suavização);
 - Valores grandes de λ produzem menor flexibilidade (maior suavização);
 - λ é estimado a partir dos dados;

Modelos aditivos generalizados

- Na verossimilhança penalizada, λ é o parâmetro de suavização, que vai controlar o nível de suavização de $s(\cdot)$.
- Em modelos com vários termos suaves, cada termo possui, em geral, seu próprio parâmetro de suavização.
- A definição do parâmetro de suavização é crucial para o ajuste do GAM:
 - Valores pequenos de λ produzem maior flexibilidade (menor suavização);
 - Valores grandes de λ produzem menor flexibilidade (maior suavização);
 - λ é estimado a partir dos dados;
 - Na biblioteca `mgcv`, uma opção usual é o método da máxima verossimilhança restrita.

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.
- Esse controle se reflete nos **graus de liberdade efetivos (EDF)** do termo suave:

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.
- Esse controle se reflete nos **graus de liberdade efetivos (EDF)** do termo suave:
 - $\lambda \downarrow \Rightarrow$ menor penalização $\Rightarrow$ maior flexibilidade $\Rightarrow$ maior EDF;

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.
- Esse controle se reflete nos **graus de liberdade efetivos (EDF)** do termo suave:
 - $\lambda \downarrow \Rightarrow$ menor penalização $\Rightarrow$ maior flexibilidade $\Rightarrow$ maior EDF;
 - $\lambda \uparrow \Rightarrow$ maior penalização $\Rightarrow$ maior suavidade $\Rightarrow$ menor EDF.

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.
- Esse controle se reflete nos **graus de liberdade efetivos (EDF)** do termo suave:
 - $\lambda \downarrow \Rightarrow$ menor penalização $\Rightarrow$ maior flexibilidade $\Rightarrow$ maior EDF;
 - $\lambda \uparrow \Rightarrow$ maior penalização $\Rightarrow$ maior suavidade $\Rightarrow$ menor EDF.
- Os **graus de liberdade efetivos (EDF)** quantificam a complexidade efetivamente utilizada pelo termo suave. Em particular, $\text{EDF} \approx 1$ indica um efeito aproximadamente linear.

Modelos aditivos generalizados

- O parâmetro λ controla o compromisso entre **ajuste aos dados** e **suavidade** da função estimada.
- Esse controle se reflete nos **graus de liberdade efetivos (EDF)** do termo suave:
 - $\lambda \downarrow \Rightarrow$ menor penalização $\Rightarrow$ maior flexibilidade $\Rightarrow$ maior EDF;
 - $\lambda \uparrow \Rightarrow$ maior penalização $\Rightarrow$ maior suavidade $\Rightarrow$ menor EDF.
- Os **graus de liberdade efetivos (EDF)** quantificam a complexidade efetivamente utilizada pelo termo suave. Em particular, $\text{EDF} \approx 1$ indica um efeito aproximadamente linear.
- A dimensão da base (k) limita a flexibilidade disponível; o EDF indica quanto dessa flexibilidade é efetivamente utilizado após a penalização.

Exemplo 1 - Emissão de gases na combustão de etanol

- Vamos retomar a análise dos dados de emissão de gases na combustão de etanol.

Exemplo 1 - Emissão de gases na combustão de etanol

- Vamos retomar a análise dos dados de emissão de gases na combustão de etanol.
- O seguinte GAM será ajustado:

$$\text{NO}_x | \mathbf{E} \sim \text{Normal}(\mu_{\mathbf{E}}, \sigma^2)$$

$$\mu_{\mathbf{E}} = \beta_0 + s(\mathbf{E})$$

Exemplo 1 - Emissão de gases na combustão de etanol

- Vamos retomar a análise dos dados de emissão de gases na combustão de etanol.
- O seguinte GAM será ajustado:

$$\text{NO}_x | \mathbf{E} \sim \text{Normal}(\mu_{\mathbf{E}}, \sigma^2)$$

$$\mu_{\mathbf{E}} = \beta_0 + s(\mathbf{E})$$

- Vamos avaliar o efeito de diferentes valores de λ sobre o EDF e a curva estimada, além de comparar diferentes bases para a função suave (os valores de λ serão fixados para fins didáticos).

Exemplo 1 - Emissão de gases na combustão de etanol

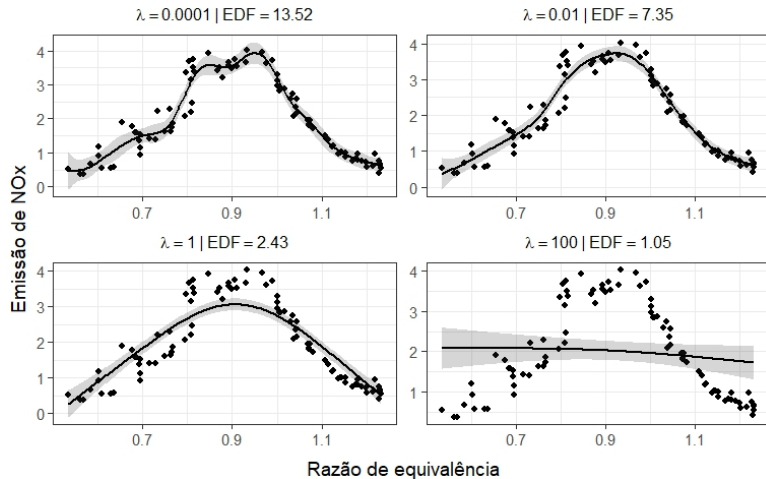


Figura 2: GAM ajustado para diferentes escolhas de λ (EDF).

Exemplo 1 - Emissão de gases na combustão de etanol

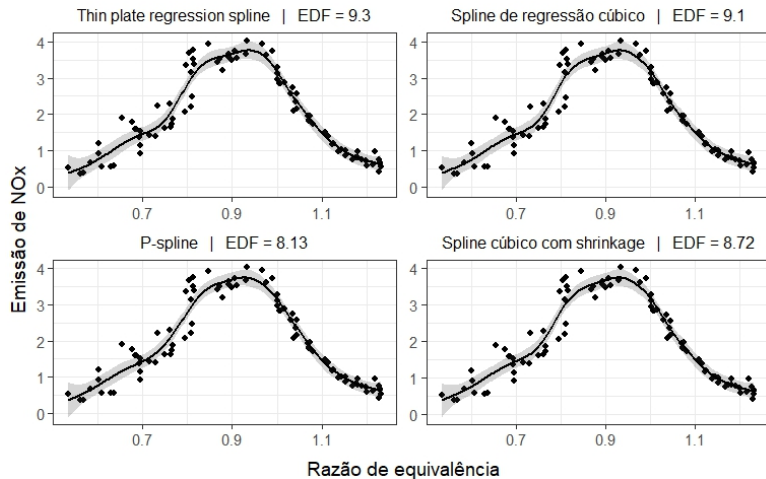


Figura 3: GAM ajustado para diferentes escolhas de função de base.

Exemplo 1 - Emissão de gases na combustão de etanol

- Com bases razoáveis, os ajustes podem ser bastante semelhantes; a penalização é determinante para a complexidade efetivamente utilizada.

Exemplo 2 - Concentração de ozônio na atmosfera

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:
 - `O3`: Concentração de ozônio (em ppm);

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:
 - `O3`: Concentração de ozônio (em ppm);
 - `ibh`: Altura da base da inversão;

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:
 - `O3`: Concentração de ozônio (em ppm);
 - `ibh`: Altura da base da inversão;
 - `ibt`: Temperatura na base da inversão;

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:
 - `O3`: Concentração de ozônio (em ppm);
 - `ibh`: Altura da base da inversão;
 - `ibt`: Temperatura na base da inversão;
 - `temp`: Temperatura.

Exemplo 2 - Concentração de ozônio na atmosfera

- Nesta aplicação vamos analisar dados atmosféricos registrados em Los Angeles, 1976. Os dados estão disponíveis na base `Ozone` da biblioteca `faraway`.
- A base é composta por $n = 330$ registros e as variáveis que vamos analisar são as seguintes:
 - `O3`: Concentração de ozônio (em ppm);
 - `ibh`: Altura da base da inversão;
 - `ibt`: Temperatura na base da inversão;
 - `temp`: Temperatura.
- O objetivo é ajustar um modelo de regressão que permita prever a concentração de ozônio com base em `ibh`, `ibt` e `temp`.

Exemplo 2 - Concentração de ozônio na atmosfera

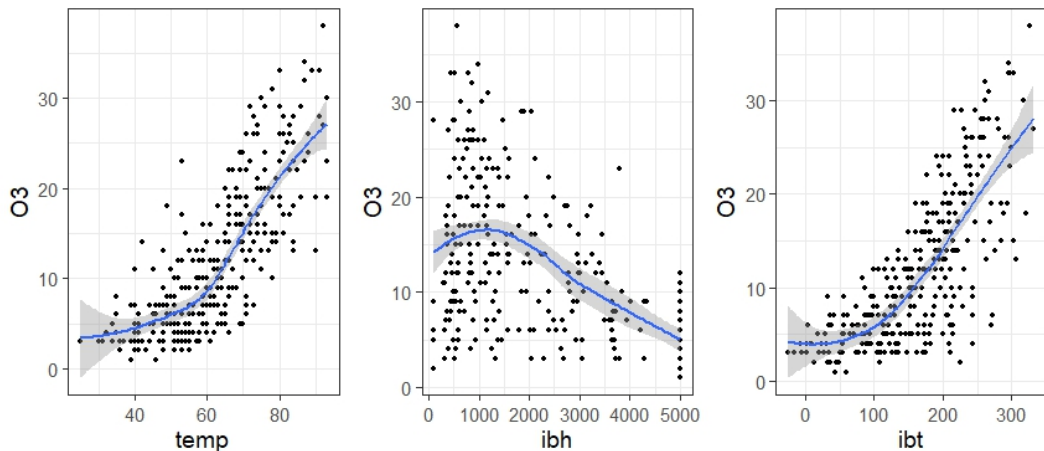


Figura 4: Relação entre a concentração de ozônio e os preditores.

Exemplo 2 - Concentração de ozônio na atmosfera

- Podemos observar relações não lineares entre a concentração de ozônio e os preditores.

Exemplo 2 - Concentração de ozônio na atmosfera

- Podemos observar relações não lineares entre a concentração de ozônio e os preditores.
- Inicialmente, foi considerado um GAM com resposta normal e suavizadores para as três variáveis:

$$O_3 | \mathbf{x} \sim \text{Normal}(\mu_x, \sigma^2)$$

$$\mu_x = \beta_0 + s_1(\text{temp}) + s_2(\text{ibh}) + s_3(\text{ibt})$$

Exemplo 2 - Concentração de ozônio na atmosfera

- Podemos observar relações não lineares entre a concentração de ozônio e os preditores.
- Inicialmente, foi considerado um GAM com resposta normal e suavizadores para as três variáveis:

$$O_3 | \mathbf{x} \sim \text{Normal}(\mu_x, \sigma^2)$$

$$\mu_x = \beta_0 + s_1(\text{temp}) + s_2(\text{ibh}) + s_3(\text{ibt})$$

- Possíveis simplificações, como substituição por termos paramétricos, são avaliadas ao longo da análise.

Exemplo 3 - Dados de cirurgia espinhal

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:
 - **Kyphosis**: fator com níveis `absent` e `present` indicando a ausência ou presença de cifose;

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:
 - **Kyphosis**: fator com níveis `absent` e `present` indicando a ausência ou presença de cifose;
 - **Age**: idade (em meses);

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:
 - **Kyphosis**: fator com níveis `absent` e `present` indicando a ausência ou presença de cifose;
 - **Age**: idade (em meses);
 - **Number**: número de vértebras envolvidas na cirurgia;

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:
 - **Kyphosis**: fator com níveis `absent` e `present` indicando a ausência ou presença de cifose;
 - **Age**: idade (em meses);
 - **Number**: número de vértebras envolvidas na cirurgia;
 - **Start**: posição da primeira vértebra operada (a partir da superior).

Exemplo 3 - Dados de cirurgia espinhal

- Nesta aplicação vamos analisar dados de cirurgias de correção espinhal em crianças. Os dados estão disponíveis na base `kyphosis` da biblioteca `rpart`.
- Os dados referem-se a 81 crianças que realizaram a cirurgia. As variáveis que vamos analisar são as seguintes:
 - **Kyphosis**: fator com níveis `absent` e `present` indicando a ausência ou presença de cifose;
 - **Age**: idade (em meses);
 - **Number**: número de vértebras envolvidas na cirurgia;
 - **Start**: posição da primeira vértebra operada (a partir da superior).
- O objetivo é ajustar um modelo de regressão para explicar a probabilidade de ocorrência de cifose em função das demais variáveis.

Exemplo 3 - Dados de cirurgia espinhal

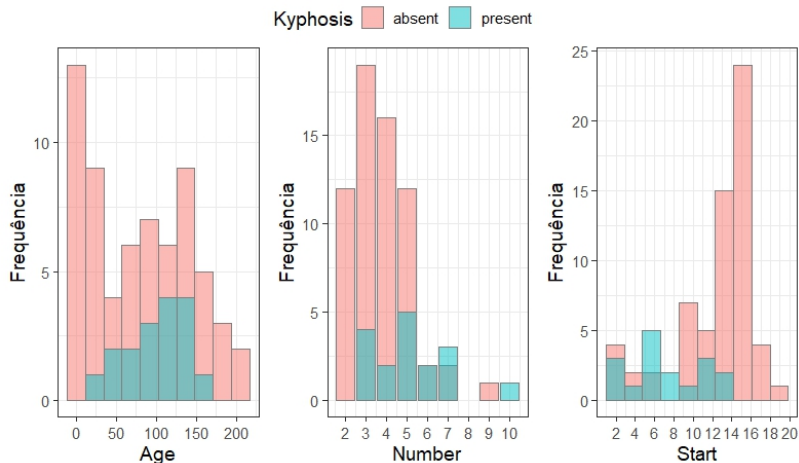


Figura 5: Presença de cifose segundo os valores das covariáveis.

Exemplo 3 - Dados de cirurgia espinhal

- Inicialmente, foi considerado um GAM com funções suaves para **Age** e **Start**:

$$\text{Kyphosis}|\mathbf{x} \sim \text{Bernoulli}(\pi_{\mathbf{x}})$$

$$\log\left(\frac{\pi_{\mathbf{x}}}{1 - \pi_{\mathbf{x}}}\right) = \beta_0 + \beta_1 \text{Number} + s_1(\text{Age}) + s_2(\text{Start})$$

Exemplo 3 - Dados de cirurgia espinhal

- Inicialmente, foi considerado um GAM com funções suaves para **Age** e **Start**:

$$\text{Kyphosis}|\mathbf{x} \sim \text{Bernoulli}(\pi_{\mathbf{x}})$$

$$\log\left(\frac{\pi_{\mathbf{x}}}{1 - \pi_{\mathbf{x}}}\right) = \beta_0 + \beta_1 \text{Number} + s_1(\text{Age}) + s_2(\text{Start})$$

- Possíveis simplificações, como substituição de suavizadores por termos paramétricos, serão avaliadas ao longo da análise.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.
- Os dados referem-se aos escores médios dos estudantes de 65 países, e as variáveis que vamos analisar são as seguintes:

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.
- Os dados referem-se aos escores médios dos estudantes de 65 países, e as variáveis que vamos analisar são as seguintes:
 - **Overall**: Escore médio registrado em Ciências para estudantes de 15 anos;

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.
- Os dados referem-se aos escores médios dos estudantes de 65 países, e as variáveis que vamos analisar são as seguintes:
 - **Overall**: Escore médio registrado em Ciências para estudantes de 15 anos;
 - **Income**: Índice de renda (numérico);

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.
- Os dados referem-se aos escores médios dos estudantes de 65 países, e as variáveis que vamos analisar são as seguintes:
 - **Overall**: Escore médio registrado em Ciências para estudantes de 15 anos;
 - **Income**: Índice de renda (numérico);
 - **Edu**: Índice de escolaridade (numérico).

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Nesta aplicação vamos analisar dados do exame PISA (Programme for International Student Assessment), coletados em 2005. A base é disponibilizada pelo professor.
- Os dados referem-se aos escores médios dos estudantes de 65 países, e as variáveis que vamos analisar são as seguintes:
 - **Overall**: Escore médio registrado em Ciências para estudantes de 15 anos;
 - **Income**: Índice de renda (numérico);
 - **Edu**: Índice de escolaridade (numérico).
- O objetivo é ajustar um modelo de regressão para explicar o escore médio em Ciências em função dos demais índices.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

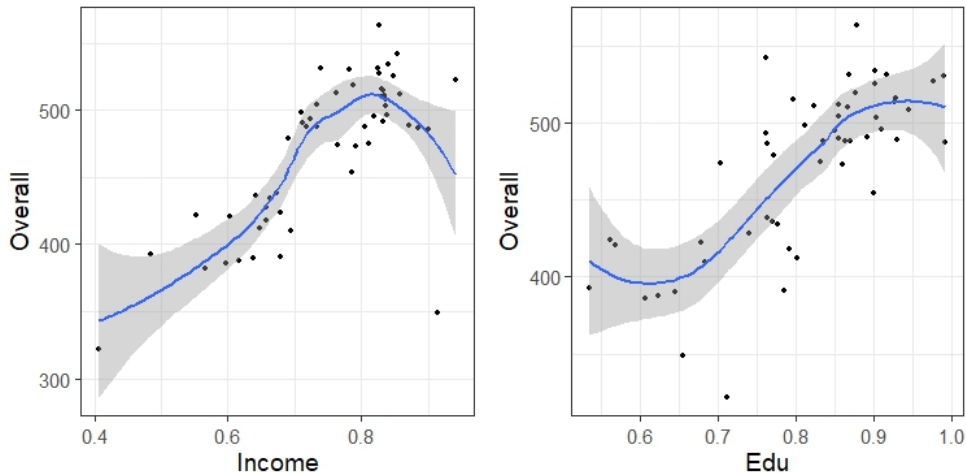


Figura 6: Relação entre o desempenho em Ciências e os índices de renda e escolaridade.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

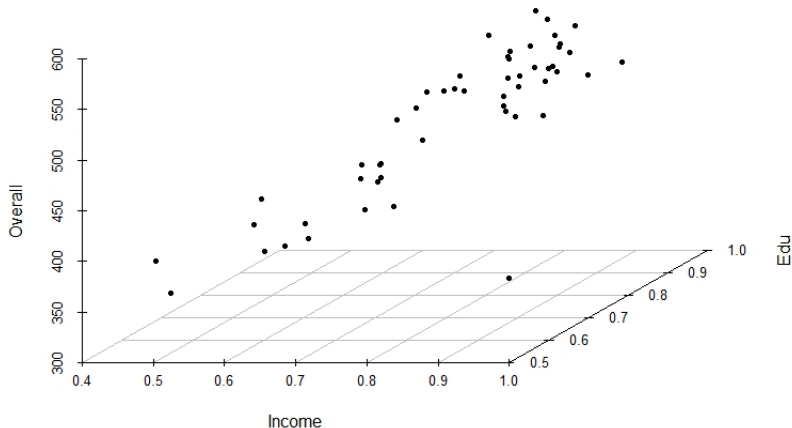


Figura 7: Representação dos dados em três dimensões.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Neste exemplo, vamos considerar um suavizador bidimensional para representar o efeito conjunto de **Income** e **Edu**, permitindo uma possível interação entre os preditores.

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Neste exemplo, vamos considerar um suavizador bidimensional para representar o efeito conjunto de **Income** e **Edu**, permitindo uma possível interação entre os preditores.
- Modelo aditivo:

$$\text{Overall}|x \sim \text{Normal}(\mu_x, \sigma^2)$$

$$\mu_x = \beta_0 + s_1(\text{Income}) + s_2(\text{Edu})$$

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Neste exemplo, vamos considerar um suavizador bidimensional para representar o efeito conjunto de **Income** e **Edu**, permitindo uma possível interação entre os preditores.
- Modelo aditivo:

$$\text{Overall}|x \sim \text{Normal}(\mu_x, \sigma^2)$$

$$\mu_x = \beta_0 + s_1(\text{Income}) + s_2(\text{Edu})$$

- Efeito conjunto:

$$\mu_x = \beta_0 + s(\text{Income}, \text{Edu})$$

Exemplo 4 - Desempenho no PISA: renda e escolaridade

- Neste exemplo, vamos considerar um suavizador bidimensional para representar o efeito conjunto de **Income** e **Edu**, permitindo uma possível interação entre os preditores.
- Modelo aditivo:

$$\text{Overall}|x \sim \text{Normal}(\mu_x, \sigma^2)$$

$$\mu_x = \beta_0 + s_1(\text{Income}) + s_2(\text{Edu})$$

- Efeito conjunto:

$$\mu_x = \beta_0 + s(\text{Income}, \text{Edu})$$

- Outros modelos também são avaliados na análise.

O que aprendemos nesta aula

O que aprendemos nesta aula

- GAMs estendem GLMs permitindo substituir alguns efeitos lineares por funções suaves quando necessário;

O que aprendemos nesta aula

- GAMs estendem GLMs permitindo substituir alguns efeitos lineares por funções suaves quando necessário;
- A penalização controla a complexidade dos termos suaves, resumida pelo EDF;

O que aprendemos nesta aula

- GAMs estendem GLMs permitindo substituir alguns efeitos lineares por funções suaves quando necessário;
- A penalização controla a complexidade dos termos suaves, resumida pelo EDF;
- Termos paramétricos e suaves podem coexistir no mesmo modelo;

O que aprendemos nesta aula

- GAMs estendem GLMs permitindo substituir alguns efeitos lineares por funções suaves quando necessário;
- A penalização controla a complexidade dos termos suaves, resumida pelo EDF;
- Termos paramétricos e suaves podem coexistir no mesmo modelo;
- Diagnóstico, dimensão da base e interpretação dos efeitos são partes essenciais do ajuste.