YNFN para seleção de variáveis preditivas numéricas

Generalização

$$X \in \mathscr{M}_{H\times n}(\mathbb{R})\,;\, h \in \{1,\cdots,H\}\,;\, v \in \{1,\cdots,n\} \tag{1}$$

$$Y \in \mathbb{R}^H, \text{ vetor ou matriz de uma coluna} \tag{2}$$

$$n_{FP} \in \mathbb{N}, \text{ hiperparâmetro}\,;\, p \equiv n_{FP}-1 \tag{3}$$

$$x^i, x^f \in \mathbb{R}^n, \text{ hiperparâmetros} \tag{4}$$

$$w \in \mathscr{M}_{n\times n_{FP}}(\mathbb{R})\,;\, w^0_{ik} = 1 \tag{5}$$

$$\gamma \in \mathbb{R}^n\,;\, \gamma_v \equiv \frac{x^f_v - x^i_v}{p} \tag{6}$$

$$h \leftarrow 1 \tag{7}$$

$$j \in \{1,2,\cdots,p\}^n\,;\, j^h_v = \left(\left\lfloor \frac{X^h_v - x^i_v}{\gamma_v} \right\rfloor + 1\right) \wedge p \vee 1 \tag{8}$$

$$_ax \in \mathbb{R}^n\,;\, {_ax^h_v} = x^i_v + (j^h_v - 1)\gamma_v \tag{9}$$

$$\mu \in \mathbb{R}^n\,;\, f({_ax^h_v}) = 1\,;\, f({_ax^h_v}+\gamma_v) = 0 \Rightarrow \mu^h_v = f(X^h_v) = \frac{0-1}{\gamma_v}X^h_v + \frac{_ax^h_v}{\gamma_v} + 1 \tag{10}$$

$$y^h_s = \sum_{v=1}^{n} [\mu^h_v w^{h-1}_{v,j^h_v} + (1-\mu^h_v) w^{h-1}_{v,j^h_v+1}] \tag{11}$$

$$\alpha^h = \frac{1}{\sum\limits_{v=1}^{n} [(\mu^h_v)^2 + (1-\mu^h_v)^2]} \tag{12}$$

$$w^h_{v,j^h_v} = w^{h-1}_{v,j^h_v} - \alpha^h (y^h_s - Y^h)\mu^h_v \tag{13}$$

$$w^h_{v,j^h_v+1} = w^{h-1}_{v,j^h_v+1} - \alpha^h (y^h_s - Y^h)(1-\mu^h_v) \tag{14}$$

$$h \leftarrow h+1 \text{ and go to } j^h \tag{15}$$

$$W \in \mathbb{R}^n\,;\, W_v = \frac{\sum\limits_{k=1}^{n_{FP}} |w^H_{vk}|}{n_{FP}} \text{ e queremos apenas os índices } v \text{ tais que } W_v > \frac{0.70}{n}. \tag{16}$$

Por outro lado, como estamos minimizando $E = 0.5\sum_h (y^h_s - Y^h)^2$, ao derivarmos em relação a w, o gradiente zero nos diz que:

$$(y^h_s - Y^h)\mu^h_v = 0 \therefore \mu^h_v = 0 \text{ (em particular) ou } y^h_s = Y^h \text{ (em geral)} \tag{17}$$

$$(y^h_s - Y^h)(1-\mu^h_v) = 0 \Rightarrow \sum_{v=1}^{n} [\mu^h_v w_{v,j_v} + (1-\mu^h_v) w_{v,j^h_v+1}] = Y = A\tilde{W} \tag{18}$$

$$\Rightarrow \tilde{W} = \underbrace{(A^\top A)^{-1}}_{m\times m} \underbrace{A^\top}_{m\times H} Y\,;\, A \in \mathscr{M}_{H\times m}(\mathbb{R})\,;\, m = n\cdot n_{FP} \tag{19}$$

$$n' < n \Rightarrow \tilde{W}' = \underbrace{(A^\top A)^{-1}}_{m'\times m'} \underbrace{A^\top}_{m'\times H} Y\,;\, A \in \mathscr{M}_{H\times m'}(\mathbb{R})\,;\, m' = n'\cdot n_{FP} \tag{20}$$

Como criar $A^h_{1\times m}$: Redimensione $p(M_{n\times n_{FP}}) = \sum_{i=1}^{n} \delta_{1,i} M \sum_{k=1}^{n_{FP}} \delta_{k,10i-10+k}$.

Como criar $W_{n\times n_{FP}}$: Redimensione $q(\tilde{W}_{m\times 1}) = \sum_{i=1}^{n}\delta_{i,1}\tilde{W}^\top\sum_{k=1}^{n_{FP}}\delta_{10k-10+k,k}$.
Suponha que $n_{FP}=2$.

$$p=j=1\,;\,w\in\mathscr{M}_{n\times 2}(\mathbb{R})\,;\,w_{ik}^0=1\,;\,\gamma_v=x_v^f-x_v^i\,;\,{}_ax_v=x_v^i \tag{21}$$

$$h\leftarrow 1 \tag{22}$$

$$f(x_v^i)=1\,;\,f(x_v^i+\gamma_v)=0\Rightarrow\mu_v^h=f(X_v^h)=\frac{x_v^i-X_v^h}{\cancel{\gamma_v}x_v^f-x_v^i}+2=AX_v^h+B \tag{23}$$

$$y_s^h(w^{h-1})=\sum_{v=1}^{n}\mu_v^h w_{v,1}^{h-1}+\sum_{v=1}^{n}(1-\mu_v^h)w_{v,2}^{h-1} \tag{24}$$

$$w_{v,1}^h=w_{v,1}^{h-1}-\frac{AX_v^h+B}{\sum_{v=1}^{n}(AX_v^h+B)^2+\sum_{v=1}^{n}(AX_v^h+B-1)^2}\left[\sum_{v=1}^{n}(AX_v^h+B)w_{v,1}^{h-1}+\sum_{v=1}^{n}(1-AX_v^h-B)w_{v,2}^{h-1}-Y^h\right] \tag{25}$$

$$w_{v,2}^h=w_{v,2}^{h-1}-\frac{AX_v^h+B}{\sum_{v=1}^{n}(AX_v^h+B)^2+\sum_{v=1}^{n}(AX_v^h+B-1)^2}\left[\sum_{v=1}^{n}(AX_v^h+B)w_{v,1}^{h-1}+\sum_{v=1}^{n}(1-AX_v^h-B)w_{v,2}^{h-1}-Y^h\right] \tag{26}$$

$$h\leftarrow h+1 \text{ and go to } w_{v,1}^h \tag{27}$$

$$W\in\mathbb{R}^n\,;\,W_v=\frac{|w_{v,1}^H|+|w_{v,2}^H|}{2}\text{ e queremos apenas os índices } v \text{ tais que } W_v>\frac{0.70}{n}. \tag{28}$$

Seja δ_{ij} uma matriz inteiramente igual a zero, exceto na linha i e na coluna j, em que é unitária.
Seja também $\varepsilon_v=\sum_{i=1}^{n}\delta_{iv}$ uma matriz que, quando aplicada a W, retorna sua v-ésima coluna.
Vamos desenvolver a linha (25), utilizando matrizes K independentes de w:

$$w_{:,1}^h=w_{:,1}^{h-1}-K_1(\langle K_2,w_{:,1}^{h-1}\rangle+\langle K_3,w_{:,2}^{h-1}\rangle-Y^h)K_2 \tag{29}$$

$$w^h\varepsilon_1=w^{h-1}\varepsilon_1-K_1(K_2^\top w^{h-1}\varepsilon_1K_2-K_3^\top w^{h-1}\varepsilon_2K_2+Y^hK_2) \tag{30}$$

Repetimos o raciocínio para a linha (26):

$$w_{:,2}^h=w_{:,2}^{h-1}-K_1(\langle K_2,w_{:,1}^{h-1}\rangle+\langle K_3,w_{:,2}^{h-1}\rangle-Y^h)K_2 \tag{31}$$

$$w^h\varepsilon_2=w^{h-1}\varepsilon_2-K_1(K_2^\top w^{h-1}\varepsilon_1K_2-K_3^\top w^{h-1}\varepsilon_2K_2+Y^hK_2) \tag{32}$$

Finalmente,

$$w^h=w^h\underbrace{\varepsilon_1}_{2\times 1}\underbrace{\delta_{11}}_{1\times 2}+w^h\varepsilon_2\delta_{12}=\sum_{i=1}^{2}[w^{h-1}\varepsilon_i+K_1(-K_2^\top w^{h-1}\varepsilon_1K_2-K_3^\top w^{h-1}\varepsilon_2K_2+Y^hK_2)]\delta_{1,i} \tag{33}$$

$$=\sum_{i=1}^{2}\sum_{\ell=1}^{3}(K_{4,\ell}w^{h-1}K_{5,i,\ell})\delta_{1,i}+K_1Y^hK_2(\delta_{11}+\delta_{12})=F(w^{h-1}) \tag{34}$$

$$w^1=F(w^0) \tag{35}$$

$$w^2=F(w^1)=F^2(w^0) \tag{36}$$

$$w^h=F^h(w^0) \tag{37}$$

Em que:

$$K_1 = \frac{1}{\sum_{v=1}^{n}\left(AX_v^h + B\right)^2 + \sum_{v=1}^{n}\left(AX_v^h + B - 1\right)^2} \in \mathbb{R} \tag{38}$$

$$K_2 = AX^h + B \in \mathbb{R}^n\,;\, K_3 = \underbrace{\varepsilon_1}_{n\times 1} - K_2 \tag{39}$$

$$K_{4,1} = I_{n\times n}\,;\, K_{4,2} = -K_1K_2^\top\,;\, K_{4,3} = -K_1K_3^\top \tag{40}$$

$$K_{5,i,1} = \varepsilon_i\,;\, K_{5,i,2} = \varepsilon_1 K_2\,;\, K_{5,i,3} = \varepsilon_2 K_2 \tag{41}$$

Para um número de funções de pertinência qualquer, são feitas as adaptações:

$$\gamma \equiv \frac{x^f - x^i}{p} \in \mathbb{R}^n \tag{42}$$

$$h \leftarrow 1 \tag{43}$$

$$j_v^h = \left(\left\lfloor \frac{X_v^h - x_v^i}{\gamma_v} \right\rfloor + 1\right) \wedge p \vee 1 \in \{1, \cdots, p\} \tag{44}$$

$$_a x_v^h = x_v^i + (j_v^h - 1)\gamma_v \in \mathbb{R} \tag{45}$$

$$A = -\frac{1}{\gamma_v}\,;\, B = \frac{_a x_v^h}{\gamma_v} + 1 \tag{46}$$

$$(30) \Rightarrow w^h \varepsilon_{j_v^h} = w^{h-1}\varepsilon_{j_v^h} + K_1(-K_2^\top w^{h-1}\varepsilon_{j_v^h}K_2 - K_3^\top w^{h-1}\varepsilon_{j_v^h+1}K_2 + Y^h K_2) \tag{47}$$

$$(33) \Rightarrow w^h = \sum_{i=1, i\neq j_v^h, j_v^h+1}^{n_{FP}} w^{h-1} \underbrace{\varepsilon_i}_{n_{FP}\times 1} \underbrace{\delta_{1,i}}_{1\times n_{FP}} + \tag{48}$$

$$+ \sum_{i=j_v^h}^{j_v^h+1} \sum_{\ell=1}^{3} (K_{4,\ell} w^{h-1} K_{5,i,\ell})\delta_{1,i} + K_1 Y^h K_2(\delta_{1,j_v^h} + \delta_{1,j_v^h+1}) = F(w^{h-1}) \tag{49}$$

$$(41) \Rightarrow K_{5,i,1} = \varepsilon_i\,;\, K_{5,i,2} = \varepsilon_{j_v^h} K_2\,;\, K_{5,i,3} = \varepsilon_{j_v^h+1} K_2 \tag{50}$$

$$RMSE^1 = \sqrt{\frac{\sum_{h=1}^{H}(y_s^h - Y^h)^2}{H}} \tag{51}$$

$$W_v^1 = \frac{\sum_{k=1}^{n_{FP}} |w_{vk}^H|}{n_{FP}} \text{ e queremos apenas os índices } v \text{ tais que } W_v^1 > \frac{0.70}{n}. \tag{52}$$

Com tais índices, temos um novo X' com mesmo H e novo $n' \leq n$, e vamos gerar novos $RMSE^2$ e $W_{v'}^2$. Comparar RMSEs 1 e 2.

Pearson

A correlação de x_i com y é dada por

$$c_v = \frac{\mathrm{cov}(x_v, y)}{\sqrt{\mathrm{var}(x_v)\,\mathrm{var}(y)}} \tag{53}$$

$$W_v^3 = \frac{c_v}{\sum_{v=1}^{n} c_v} \text{ e queremos apenas os índices } v \text{ tais que } W_v^3 > \frac{0.70}{n}. \tag{54}$$

Com tais índices, temos um novo X'' com mesmo H e novo $n'' \leq n$, e vamos gerar novos $RMSE^3$ e $W_{v''}^4$. Podemos comparar RMSEs 2 e 3.

Podemos fixar uma variável $v = 1$ e comparar peso de Pearson com peso de YNFN: $W_v^1 \neq W_v^3$.

Introdução

$$n_{FP} = 2\,;\, x_i = 1\,;\, x_f = 3\,;\, \gamma = 2\,;\, w^0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}\,;\, x_a = 1 - \gamma \tag{55}$$

$$f(x_1, x_2, x_3) = 1 + x_1^{0.5} + x_2^{-1} + x_3^{-1.5} \tag{56}$$

$$\mu = \frac{1}{2}(3 - x_v) = \begin{pmatrix} 1 \to 1 \\ 2 \to 0.5 \\ 3 \to 0 \end{pmatrix} \tag{57}$$

$$f(1,1,1) = 4\,;\, \mu^1 = (1,1,1) \tag{58}$$

$$f(1,1,2) = 3 + \frac{\sqrt{2}}{4};\, \mu^2 = (1,1,0.5) \tag{59}$$

$$f(1,1,3) = 3 + \frac{\sqrt{3}}{9};\, \mu^3 = (1,1,0) \tag{60}$$

$$f(1,2,1) = 3.5\,;\, \mu^4 = (1,0.5,1) \tag{61}$$

$$f(1,2,2) = 2.5 + \frac{\sqrt{2}}{4};\, \mu^5 = (1,0.5,0.5) \tag{62}$$

$$f(1,2,3) = 2.5 + \frac{\sqrt{3}}{9};\, \mu^6 = (1,0.5,0) \tag{63}$$

$$f(1,3,1) = \frac{4}{3};\, \mu^7 = (1,0,1) \tag{64}$$

$$f(1,3,2) = \frac{7}{3} + \frac{\sqrt{2}}{4};\, \mu^8 = (1,0,0.5) \tag{65}$$

$$f(1,3,3) = \frac{7}{3} + \frac{\sqrt{3}}{9};\, \mu^9 = (1,0,0) \tag{66}$$

$$f(2,1,1) = 3 + \sqrt{2};\, \mu^{10} = (0.5,1,1) \tag{67}$$

$$f(2,1,2) = 2 + 1.25\sqrt{2};\, \mu^{11} = (0.5,1,0.5) \tag{68}$$

$$f(2,1,3) = 2 + \sqrt{2} + \frac{\sqrt{3}}{9};\, \mu^{12} = (0.5,1,0) \tag{69}$$

$$f(2,2,1) = 2.5 + \sqrt{2};\, \mu^{13} = (0.5,0.5,0) \tag{70}$$

$$f(2,2,2) = 1.5 + 1.25\sqrt{2};\, \mu^{14} = (0.5,0.5,0.5) \tag{71}$$

$$f(2,2,3) = 1.5 + \sqrt{2} + \frac{\sqrt{3}}{9};\, \mu^{15} = (0.5,0.5,0) \tag{72}$$

$$f(2,3,1) = \frac{7}{3} + \sqrt{2};\, \mu^{16} = (0.5,0,1) \tag{73}$$

$$f(2,3,2) = \frac{4}{3} + 1.25\sqrt{2};\, \mu^{17} = (0.5,0,0.5) \tag{74}$$

$$f(2,3,3) = \frac{4}{3} + \sqrt{2} + \frac{\sqrt{3}}{9};\, \mu^{18} = (0.5,0,0) \tag{75}$$

$$f(3,1,1) = 3 + \sqrt{3};\, \mu^{19} = (0,1,1) \tag{76}$$

$$f(3,1,2) = 2 + \sqrt{3} + \frac{\sqrt{2}}{4};\, \mu^{20} = (0,1,0.5) \tag{77}$$

$$f(3,1,3) = 2 + 10/9\sqrt{3};\, \mu^{21} = (0,1,0) \tag{78}$$

$$f(3,2,1) = 2.5 + \sqrt{3};\, \mu^{22} = (0,0.5,1) \tag{79}$$

$$f(3,2,2) = 1.5 + \sqrt{3} + \frac{\sqrt{2}}{4};\, \mu^{23} = (0,0.5,0.5) \tag{80}$$

$$f(3,2,3) = 1.5 + 10/9\sqrt{3};\, \mu^{24} = (0,0.5,0) \tag{81}$$

$$f(3,3,1) = \frac{7}{3} + \sqrt{3};\, \mu^{25} = (0,0,1) \tag{82}$$

$$f(3,3,2) = \frac{4}{3} + \sqrt{3} + \frac{\sqrt{2}}{4};\, \mu^{26} = (0,0,0.5) \tag{83}$$

$$f(3,3,3) = \frac{4}{3} + 10/9\sqrt{3};\, \mu^{27} = (0,0,0) \tag{84}$$

$$A = [\mu; 1-\mu]_{27\times 6} \tag{85}$$

$$A \cdot \tilde{W}_{6\times 1} = F^2 = Y_{27\times 1} \therefore \tilde{W} = (A^\top A)^{-1} A^\top Y \tag{86}$$

Versão de 17/abril/2023* por Vinicius Claudino Ferraz.

*Fora da caridade não há salvação.