Aprendizagem Automática

João Paulo Pordeus Gomes

Regressão Linear Univariada

Regressão Linear

- Explicar uma determinada variável de saída (y) através de variáveis de entrada (x), utilizando um modelo linear.

Regressão Linear

- Explicar uma determinada variável de saída (y) através de variáveis de entrada (x), utilizando um modelo linear.
- Exemplo
  - Concessão de crédito

Regressão Linear

- Concessão de Crédito

Salário (R$)	Crédito (R$)
1500	16500
2000	18000
3000	28000
3300	33000
4200	49000
...	...
5500	52000

Regressão Linear

- Concessão de Crédito
  - Encontrar uma relação linear entre Salário e Credito

Salário (R$)	Crédito (R$)
1500	16500
2000	18000
3000	28000
3300	33000
4200	49000
...	...
5500	52000

Regressão Linear

- Concessão de Crédito
  - Encontrar uma relação linear entre Salário e Credito
  - Salário (x_i) e Crédito (y_i)

Salário (R$)	Crédito (R$)
1500	16500
2000	18000
3000	28000
3300	33000
4200	49000
...	...
5500	52000

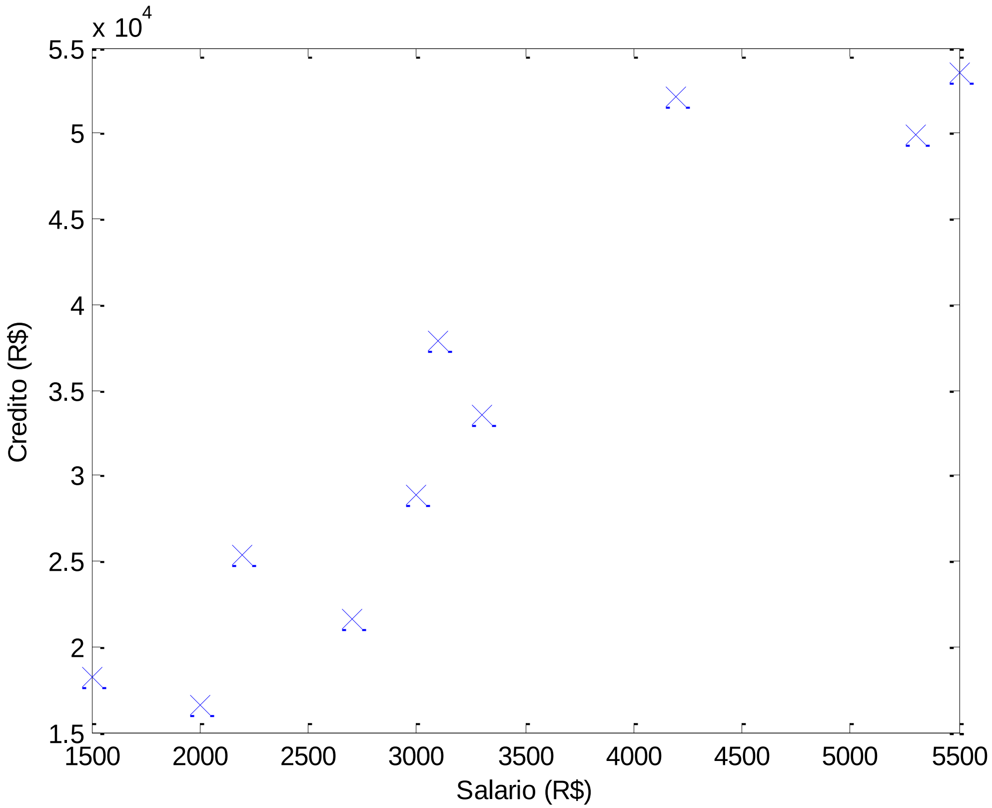
Regressão Linear

- Concessão de Crédito
  - Encontrar uma relação linear entre Salário e Credito
  - Salário (x_i) e Crédito (y_i)
  - $\bar{y}_i = w_1 x_i + w_0$

Salário (R$)	Crédito (R$)
1500	16500
2000	18000
3000	28000
3300	33000
4200	49000
...	...
5500	52000

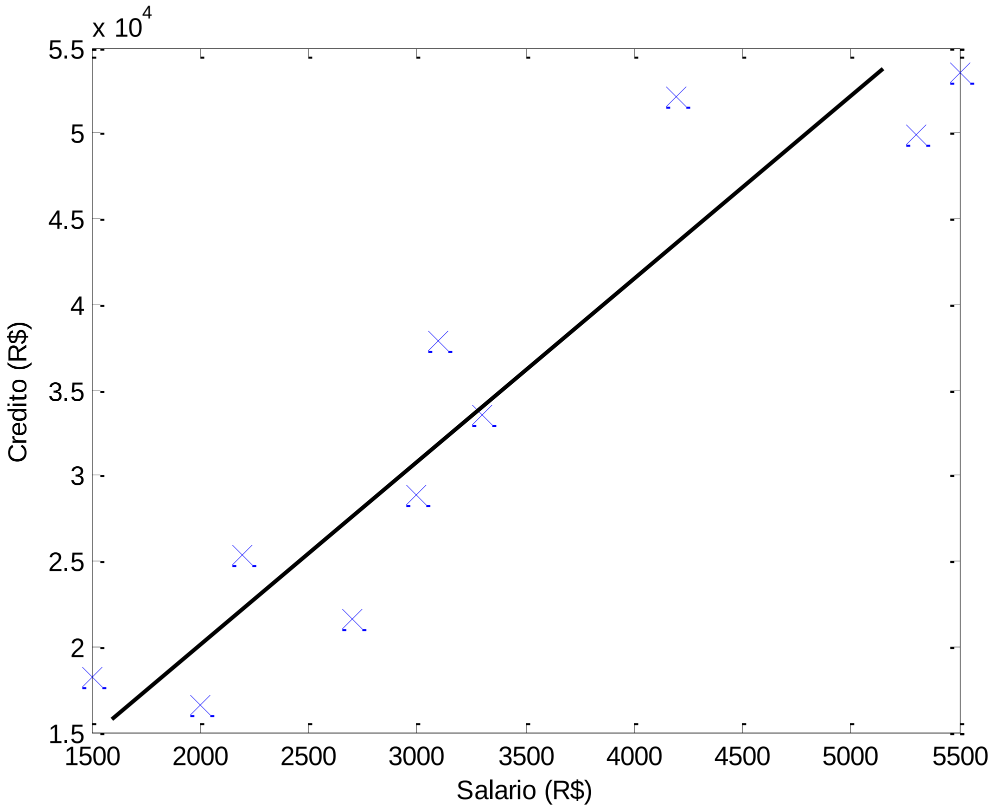
Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$



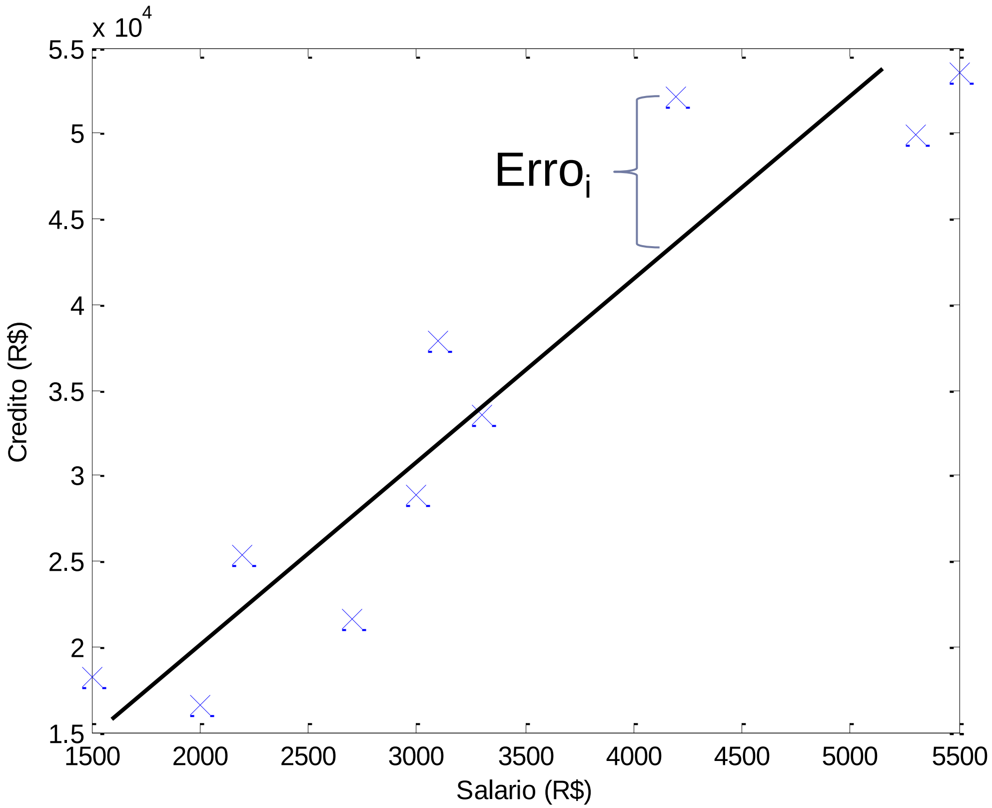
Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$



Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$



Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} e_i{}^2$
  - Onde: $e_i = y_i - \bar{y}_i$

Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} {e_i}^2$
  - Onde: $e_i = y_i - \bar{y}_i$

- $min_{w_1 w_0} J(w_1, w_0)$

Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} {e_i}^2$
  - Onde: $e_i = y_i - \bar{y}_i$

- $min_{w_1 w_0} J(w_1, w_0)$
  - $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \bar{y}_i)^2$

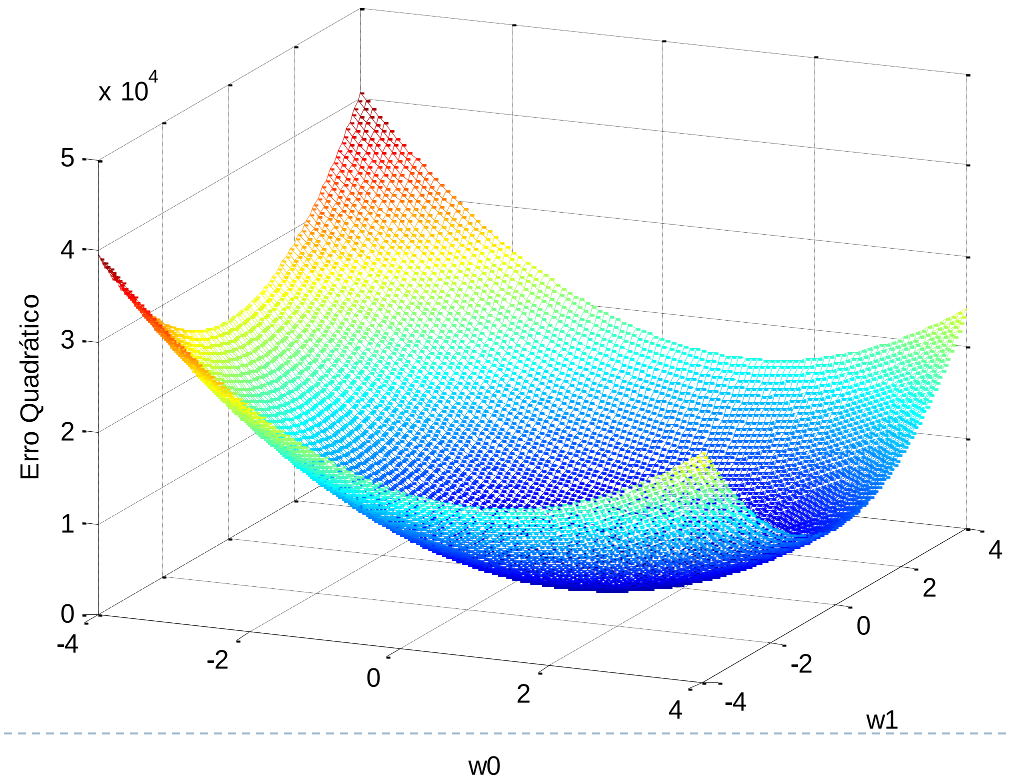
Regressão Linear

- $\bar{y}_i = w_1 x_i + w_0$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} {e_i}^2$
  - Onde: $e_i = y_i - \bar{y}_i$

- $min_{w_1 w_0} \, J(w_1, w_0)$
  - $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \bar{y}_i)^2$
  - $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

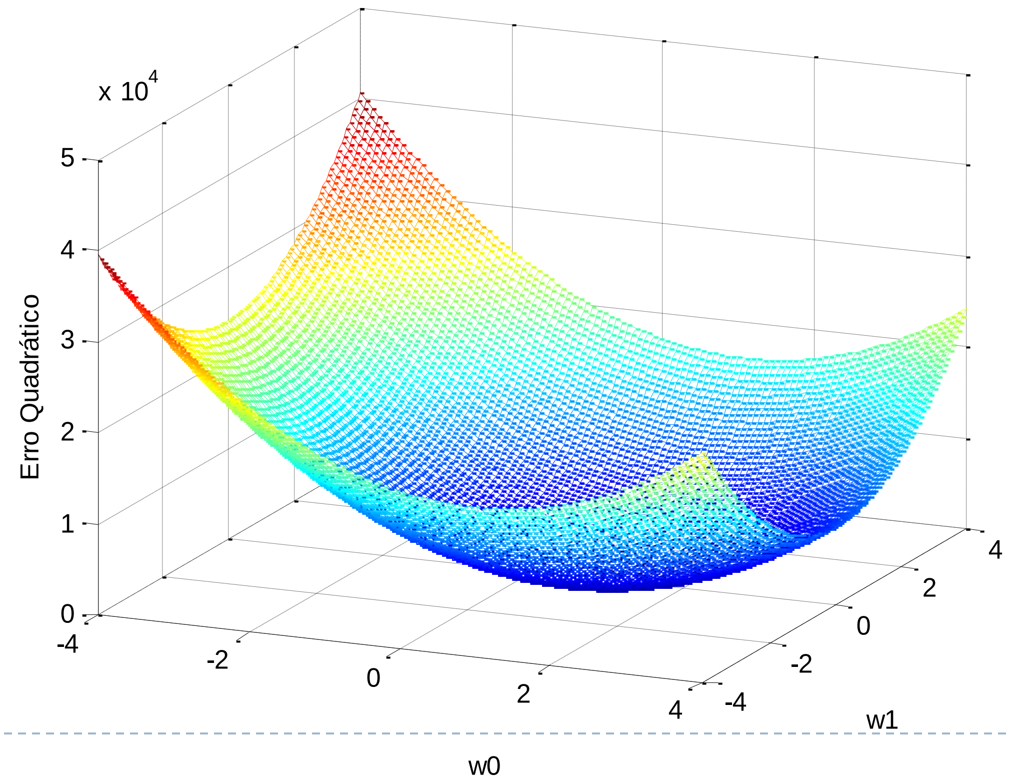
Regressão Linear

- $J(w_1, w_0) = \frac{1}{2n}\sum_{i=1}^{n}(y_i - w_1 x_i - w_0)^2$



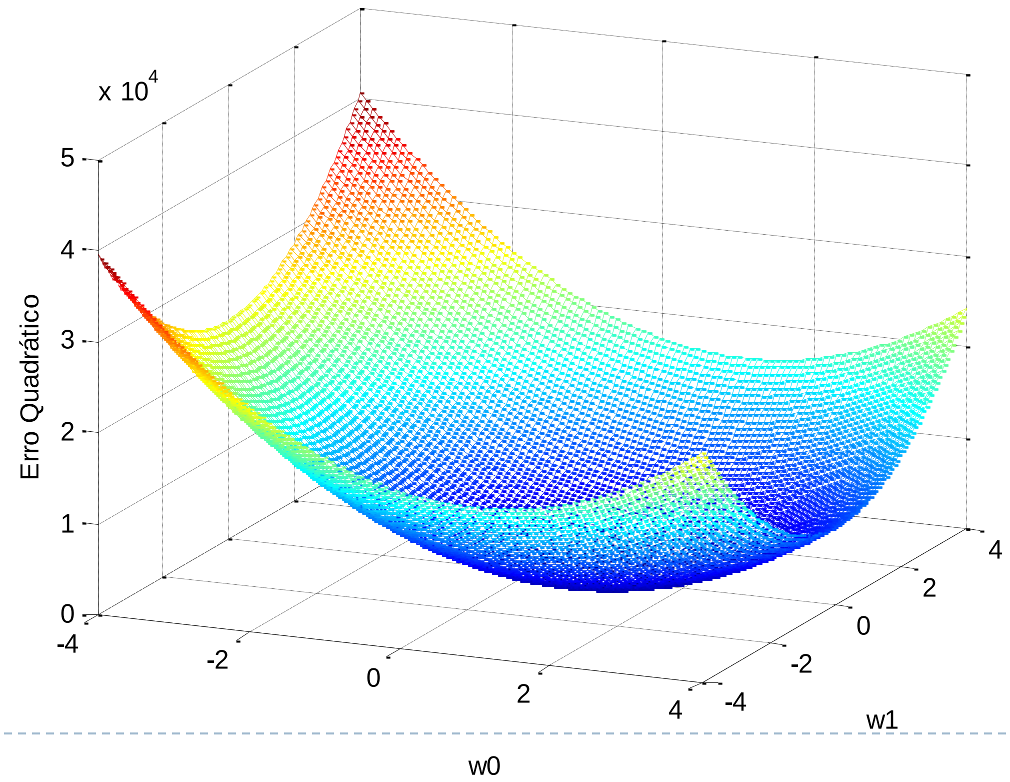
Regressão Linear

- $J(w_1, w_0) = \frac{1}{2n}\sum_{i=1}^{n}(y_i - w_1x_i - w_0)^2$
  - Método do gradiente descendente



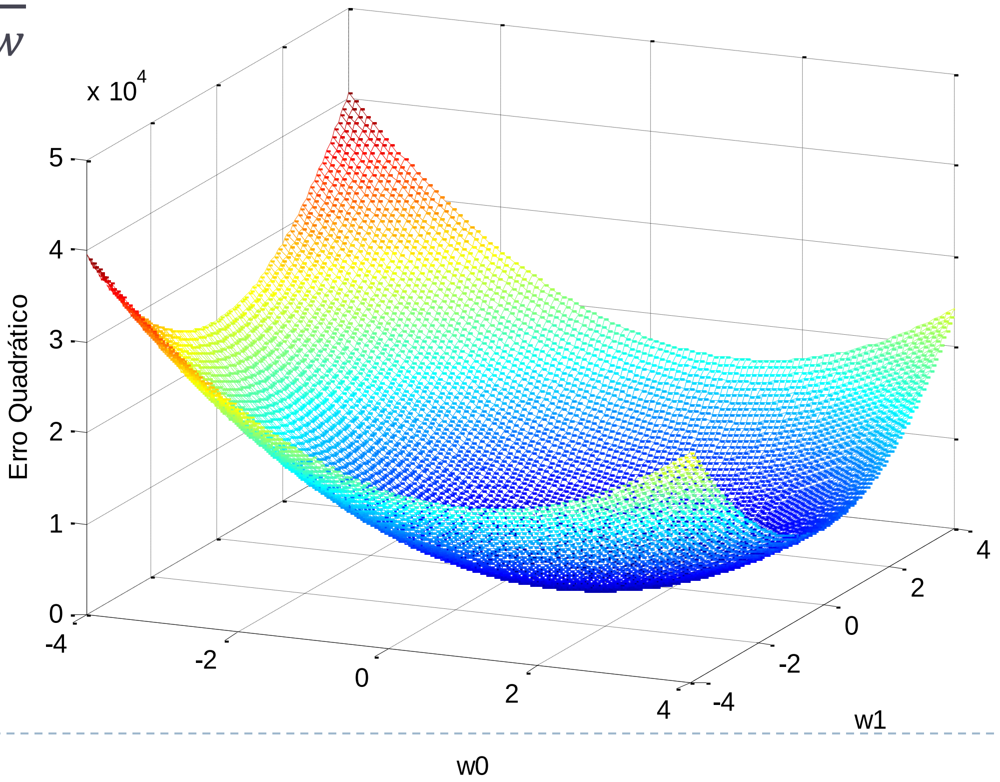
Gradiente Descendente

- Inicializar os pontos de forma aleatória
  - w_1, w_0
- Movimento os pontos para a direção que diminui J



Gradiente Descendente

- Inicializar os pontos de forma aleatória
  - w_1, w_0
- Movimento os pontos para a direção que diminui J
  - $w = w - \alpha \frac{\partial J}{\partial w}$



Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = ?$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)](-1)$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
- $\frac{\partial J}{\partial w_1} =?$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
- $\frac{\partial J}{\partial w_1} = [\frac{1}{2n} 2 \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)](-x_i)$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
- $\frac{\partial J}{\partial w_1} = (\frac{1}{2n} 2 \sum_{i=1}^{n} e_i)(-x_i)$

Gradiente Descendente

- $w = w - \alpha \frac{\partial J}{\partial w}$
- $J(w_1, w_0) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2$

- $\frac{\partial J}{\partial w_0} = [\frac{1}{2n} 2 \sum_{i=1}^{n} e_i](-1)$
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
- $\frac{\partial J}{\partial w_1} = (\frac{1}{2n} 2 \sum_{i=1}^{n} e_i)(-x_i)$
  - $w_1 = w_1 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i \, x_i$

Gradiente Descendente

- Regressão Linear Univariada
- $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
- $w_1 = w_1 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i \, x_i$

Gradiente Descendente

- Define α pequeno
- Utiliza todo o conjunto de dados e atualiza os pesos
  - $w_0 = w_0 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i$
  - $w_1 = w_1 + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i \, x_i$
- Repete o procedimento diversas vezes (épocas)

Gradiente Descendente Estocástico

- Atualização a cada amostra

Gradiente Descendente Estocástico

- Atualização a cada amostra
  - $w_0 = w_0 + \alpha e_i$
  - $w_1 = w_1 + \alpha e_i x_i$

Gradiente Descendente Estocástico

- Define α pequeno
- Utiliza todo o conjunto de dados e atualiza os pesos
  - $w_0 = w_0 + \alpha e_i$
  - $w_1 = w_1 + \alpha e_i x_i$
- Faz permutação nos dados
- Repete o procedimento diversas vezes (épocas)

Regressão Linear Multivariada

Regressão Linear Multivariada

- Em diversos problemas é necessário utilizar mais de uma variável x para tentar explicar a variável de saída y

Regressão Linear Multivariada

- Em diversos problemas é necessário utilizar mais de uma variável x para tentar explicar a variável de saída y
  - Exemplo
    - Concessão de crédito

Regressão Linear Multivariada

- Em diversos problemas é necessário utilizar mais de uma variável x para tentar explicar a variável de saída y
  - Exemplo
    - Concessão de crédito

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000
...		...
5500	1700	52000

Regressão Linear Multivariada

- Concessão de Crédito
  - Encontrar uma relação linear entre Salário,Dívida e Credito
  - Salário (x_{1i}), Salário (x_{2i}) e Crédito (y_i)
  - $\overline{y_i} = w_1 x_{1i} + w_2 x_{2i} + w_0$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000
...		...
5500	1700	52000

Regressão Linear Multivariada

- $\overline{y}_i = w_1 x_{1i} + w_2 x_{2i} + w_0$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear Multivariada

- $\overline{y}_i = w_1 x_{1i} + w_2 x_{2i} + w_0$
- $\overline{y}_i = w_1 x_{1i} + w_2 x_{2i} + w_0 x_0$
  - $x_0 = 1$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear Multivariada

- $\bar{y}_i = w_1 x_{1i} + w_2 x_{2i} + w_0$
- $\bar{y}_i = w_1 x_{1i} + w_2 x_{2i} + w_0 x_0$
  - $x_0 = 1$
- $\bar{y}_i = \boldsymbol{w}^T \boldsymbol{x}_i$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear Multivariada

- $\bar{y}_i = \boldsymbol{w}^T \boldsymbol{x}_i$
  - $\boldsymbol{x_1} = [1 \;\; 1500 \;\; 0]^T$, $\boldsymbol{x_2} = [1 \;\; 2000 \;\; 4000]^T$, ...
  - $\boldsymbol{w}^T = [w_0 \;\; w_1 \;\; w_2]$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear Multivariada

- $\bar{y}_i = \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{x}_{\boldsymbol{i}}$
- Para aprender os pesos, pode-se utiliza o gradiente descendente de forma semelhante à apresentada anteriormente.

Gradiente Descendente

- Regressão Linear Multivariada
  - $\overline{y_i} = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{x}_{\boldsymbol{i}}$
- Regra de Aprendizado
  - $\mathbf{w} = \mathbf{w} + \alpha \frac{1}{n} \sum_{i=1}^{n} e_i \, \boldsymbol{x}_{\boldsymbol{i}}$

Gradiente Descendente Estocástico

- Regressão Linear Multivariada
  - $\overline{y_i} = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{x}_{\boldsymbol{i}}$
- Regra de Aprendizado
  - $\mathbf{w} = \mathbf{w} + \alpha e_i \boldsymbol{x}_{\boldsymbol{i}}$

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.
- Exemplo
  - $\overline{y_i} = \boldsymbol{w}^T \boldsymbol{x_i}$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.
- Exemplo
  - $\overline{y_i} = \boldsymbol{w}^T \boldsymbol{x_i}$
  - $\boldsymbol{Y} = [y_1\ y_2\ \ \dots\ y_n]^T$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.
- Exemplo
  - $\overline{y_i} = \boldsymbol{w}^T \boldsymbol{x_i}$
  - $\boldsymbol{Y} = [y_1\ y_2\ \ \dots y_n]^T$
    - $\boldsymbol{Y} = [16500\ \ 12000\ \ 2800]^T$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.
- Exemplo
  - $\overline{y_i} = \boldsymbol{w}^T \boldsymbol{x_i}$
  - $\boldsymbol{Y} = [y_1 \; y_2 \;\; \dots \; y_n]^T$
  - $\boldsymbol{X} = [\boldsymbol{x_1} \;\; \boldsymbol{x_2} \;\; \dots \boldsymbol{x_n}]^T$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear (em lote - *batch*)

- Se já tivermos todos os pontos, é possível calcular os parâmetros sem utilizar um processo iterativo.
- Exemplo
  - $\overline{y_i} = \boldsymbol{w}^T \boldsymbol{x_i}$
  - $\boldsymbol{Y} = [y_1 \; y_2 \; \ldots \; y_n]^T$
  - $\boldsymbol{X} = [\boldsymbol{x_1} \; \boldsymbol{x_2} \; \ldots \boldsymbol{x_n}]^T$
    - $\boldsymbol{X} = \begin{bmatrix} 1 & 1500 & 0 \\ 1 & 2000 & 4000 \\ 1 & 3000 & 2000 \end{bmatrix}$

Salário (R$)	Dívida (R$)	Crédito (R$)
1500	0	16500
2000	4000	12000
3000	2000	28000

Regressão Linear (em lote - *batch*)

- Modelo
  - $\overline{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{w}$

Regressão Linear (em lote - *batch*)

- Modelo
  - $\overline{\boldsymbol{Y}} = \boldsymbol{Xw}$
- Função de custo
  - $J(\boldsymbol{w}\,) = \frac{1}{2}(Y - \overline{\boldsymbol{Y}})^{\boldsymbol{T}}(Y - \overline{\boldsymbol{Y}})$
  - $J(\boldsymbol{w}\,) = \frac{1}{2}(Y - \boldsymbol{Xw})^{\boldsymbol{T}}(Y - \boldsymbol{Xw})$

Regressão Linear (em lote - *batch*)

- Modelo
  - $\overline{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{w}$
- Função de custo
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \overline{\boldsymbol{Y}})^{\boldsymbol{T}}(Y - \overline{\boldsymbol{Y}})$
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \boldsymbol{X}\boldsymbol{w})^{\boldsymbol{T}}(Y - \boldsymbol{X}\boldsymbol{w})$
- Minimizando J (derivando e igualando a zero)

Regressão Linear (em lote - *batch*)

- Modelo
  - $\overline{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{w}$
- Função de custo
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \overline{\boldsymbol{Y}})^{\boldsymbol{T}}(Y - \overline{\boldsymbol{Y}})$
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \boldsymbol{X}\boldsymbol{w})^{\boldsymbol{T}}(Y - \boldsymbol{X}\boldsymbol{w})$
- Minimizando J (derivando e igualando a zero)
  - $\frac{\partial J}{\partial \boldsymbol{w}} = \frac{1}{2}\,2(-\boldsymbol{X}^{T})(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{w}) = 0$

Regressão Linear (em lote - *batch*)

- Modelo
  - $\overline{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{w}$
- Função de custo
  - $J(\boldsymbol{w}\,) = \frac{1}{2}(Y - \overline{\boldsymbol{Y}})^{\boldsymbol{T}}(Y - \overline{\boldsymbol{Y}})$
  - $J(\boldsymbol{w}\,) = \frac{1}{2}(Y - \boldsymbol{X}\boldsymbol{w})^{\boldsymbol{T}}(Y - \boldsymbol{X}\boldsymbol{w})$
- Minimizando J (derivando e igualando a zero)
  - $\frac{\partial J}{\partial \boldsymbol{w}} = \frac{1}{2}\; 2(-\boldsymbol{X}^{T})(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{w}) = 0$
  - $\boldsymbol{X}^{T}\boldsymbol{X}\boldsymbol{w} = \boldsymbol{X}^{T}\boldsymbol{Y}$

Regressão Linear (em lote - *batch*)

- Modelo
  - $\bar{\boldsymbol{Y}} = \boldsymbol{Xw}$
- Função de custo
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \bar{\boldsymbol{Y}})^{\boldsymbol{T}}(Y - \bar{\boldsymbol{Y}})$
  - $J(\boldsymbol{w}) = \frac{1}{2}(Y - \boldsymbol{Xw})^{\boldsymbol{T}}(Y - \boldsymbol{Xw})$
- Minimizando J (derivando e igualando a zero)
  - $\frac{\partial J}{\partial \boldsymbol{w}} = \frac{1}{2}\,2(-\boldsymbol{X}^{T})(\boldsymbol{Y} - \boldsymbol{Xw}) = 0$
  - $\boldsymbol{X}^{T}\boldsymbol{Xw} = \boldsymbol{X}^{T}\boldsymbol{Y}$
  - $\boldsymbol{w} = (\boldsymbol{X}^{T}\boldsymbol{X})^{-\boldsymbol{1}}\boldsymbol{X}^{T}\boldsymbol{Y}$

Regressão Linear (em lote - *batch*)

- Modelo
  - $\bar{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{w}$
- Regra de Aprendizado
  - $\boldsymbol{w} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$

Método dos mínimos quadrados

Dúvidas ?