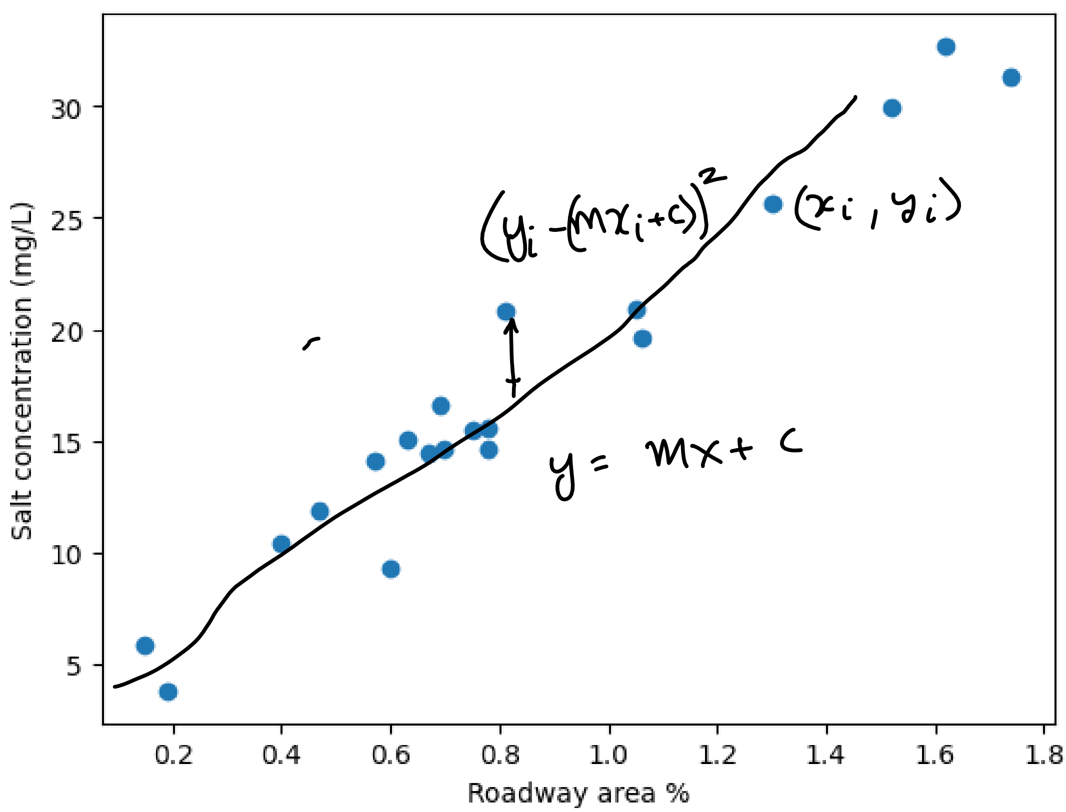
LinearModels2

February 2, 2023

1 Least Square Regression

[8]:
```
import matplotlib.pyplot as plt
# Plot the points
fig, ax = plt.subplots()
# Scatter plot using matplotlib
def plot_salt_scatter(ax, salt_concentration_data):
    ax.scatter(salt_concentration_data[:, 2], salt_concentration_data[:, 1])
    ax.set_xlabel(r"Roadway area %")
    ax.set_ylabel(r"Salt concentration (mg/L)")
plot_salt_scatter(ax, salt_concentration_data)
```



1.1 Vectorized least square

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$m^*, c^* = \arg\min_{m,c} \left\| \underline{y} - \underbrace{\left(m\,\underline{x} + c\,\mathbb{1}_n \right)} \right\|^2 \qquad (L\text{-}2 \text{ norm})$$

$$X = \underbrace{\begin{bmatrix} \underline{x} & \mathbb{1}_n \end{bmatrix}} \begin{bmatrix} m \\ c \end{bmatrix}$$

1.2 Matrix-vector product

$$\underset{\in \mathbb{R}^{n\times m}}{\underline{V}}\;\underset{\in \mathbb{R}^{m}}{\underline{u}} = \begin{bmatrix} \underline{v}_1^T \\ \underline{v}_2^T \\ \vdots \\ \underline{v}_n^T \end{bmatrix} \underline{u} = \begin{bmatrix} \underline{v}_1^T \underline{u} \\ \underline{v}_2^T \underline{u} \\ \vdots \\ \underline{v}_n^T \underline{u} \end{bmatrix} \in \mathbb{R}^n$$

$$\underline{V}\,\underline{u} = \begin{bmatrix} {}_1\underline{v} & {}_2\underline{v} & \cdots & {}_m\underline{v} \end{bmatrix} \underline{u} = \begin{bmatrix} {}_1\underline{v} & {}_2\underline{v} & \cdots & {}_m\underline{v} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}$$

$$= u_1\,{}_1\underline{v} + u_2\,{}_2\underline{v} + \cdots + u_m\,{}_m\underline{v} \in \mathbb{R}^n$$

1.3 Least square using matrices

$$\arg\min_{\underline{m}} \left\| \underline{y} - \underset{n\times 2}{X}\,\underset{2\times 1}{\underline{m}} \right\|^2, \quad \underline{m} = \begin{bmatrix} m \\ c \end{bmatrix} \qquad \underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$$

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{bmatrix} \in \mathbb{R}^{n\times 2}$$

vector-vector product

1.4 Dot product as ~~matrix~~ product

$$\underline{a} \cdot \underline{b} = a_1 b_1 + a_2 b_2 + \dots a_n b_n$$

$$\underline{a}^T \underline{b} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

$$\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a} = \underbrace{\underline{a}^T \underline{b}}_{\text{inner product}} = \underline{b}^T \underline{a}$$

$$\underbrace{a b^T}_{\text{outer product}} \neq \underbrace{a^T b}_{\text{inner product}}$$

1.5 Matrix transpose properties

1. $(A+B)^\top = ?$ $A^T + B^T$
2. $(AB)^\top = ?$ $B^T A^T$

$(ABC)^T = C^T B^T A^T$

$$\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right)^T = \begin{bmatrix} 1+2 \\ 3+3 \end{bmatrix}^T = \begin{bmatrix} 1+2 & 3+3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 3 \end{bmatrix}^T + \begin{bmatrix} 2 \\ 3 \end{bmatrix}^T$$

1.6 In-class exercise

Expand

$$\arg\min_{\underline{m}} \|\underline{y} - X\underline{m}\|^2 = \quad \in \mathbb{R}$$

$$(\mathbf{y} - \mathbf{Xm})^\top (\mathbf{y} - \mathbf{Xm}) \quad \in \mathbb{R}$$

$$(\underline{y}^T - \underline{m}^T X^T)(\underline{y} - X\underline{m}) \quad \in \mathbb{R}$$

$$= \underbrace{\underline{y}^T \underline{y}}_{\in \mathbb{R}} + \underbrace{\underline{m}^T X^T X \underline{m}}_{\in \mathbb{R}} - \underline{y}^T X \underline{m} - \underline{m}^T X^T \underline{y}$$

$$[a]^T = [a]$$

$$(\underline{y}^T X \underline{m})^T = \underline{m}^T X^T \underline{y}$$

1.7 Quadratic form

1. Single variable
2. Two variable
3. n-variable vectorized

$f(x) = ax^2 + bx + c$

$f(x,y) = ax^2 + by^2 + cxy + dx + ey + g$

$$= \underline{y}^T \underline{y} + \underline{m}^T X^T X \underline{m} - 2\underline{y}^T X \underline{m}.$$

$$= \underline{m}^T A \underline{m} + \underline{b}^T \underline{m} + d$$

$A = X^T X$
$b^T = -2y^T X$
$d = y^T y$

$$\underline{m}^T A \underline{m}, \quad m = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\underline{m}^T A \underline{m} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= a_{11} x^2 + (a_{21} + a_{12}) xy + a_{22} y^2$$

$$\underline{m}^T A \underline{m} = a_{11} m_1^2 + a_{22} m_2^2 + \ldots + a_{nn} m_n^2 + \sum_{ij} (a_{ij} + a_{ji}) m_i m_j$$

$$f(x,y) = ax^2 + by^2 + cxy + dx + ey + g$$

$$= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + g$$

$$= \underline{m}^T A \underline{m} + \underline{b}^T \underline{m} + g$$

$$\sum_i (y_i - (mx_i + c))^2$$

Jacobian

1.8 Vector derivatives

$f(\underline{x}) : \mathbb{R}^n \mapsto \mathbb{R}$

$$\frac{\partial f(\underline{x})}{\partial \underline{x}} = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(\underline{x})}{\partial x_2}, \ldots \frac{\partial f(\underline{x})}{\partial x_n} \right] = \nabla_{\underline{x}}^T f(\underline{x}) \in \mathbb{R}^n$$

$$\underline{f}(\underline{x}) : \mathbb{R}^n \mapsto \mathbb{R}^m = \underline{f}(\underline{x}) = \begin{bmatrix} f_1(\underline{x}) \\ f_2(\underline{x}) \\ \vdots \\ f_m(\underline{x}) \end{bmatrix} \in \mathbb{R}^m$$

$$\frac{\partial \underline{f}(x)}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \ldots & \frac{\partial f_1(\underline{x})}{} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & & \frac{\partial f_m(\underline{x})}{\partial x_n} \end{bmatrix} = J_{\underline{x}} \underline{f}(\underline{x}) \in \mathbb{R}^{m \times n}$$

1.9 In-class exercises

1. Find the derivative of $\mathbf{x}^\top \mathbf{A}\mathbf{x}$ with respect to $\mathbf{x}$.
2. Find the derivative of $\mathbf{b}^\top \mathbf{x}$ with respect to $\mathbf{x}$.

$\underline{x} \in \mathbb{R}^2 \quad \underline{x} = \begin{bmatrix} x \\ y \end{bmatrix} \; ; \; A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$$\frac{\partial}{\partial \underline{x}} f(\underline{x}) = \underline{x}^T A \underline{x} = a_{11}x^2 + a_{22}y^2 + (a_{12}+a_{21})xy$$

$$\frac{\partial f(\underline{x})}{\partial \underline{x}} = \left[\frac{\partial f(x)}{\partial x} \quad \frac{\partial f(\underline{x})}{\partial y} \right] = \nabla_{\underline{x}}^T f(x) = \left[2a_{11}x + (a_{12}+a_{21})y, \quad 2a_{22}y + (a_{12}+a_{21})x \right]$$

$$= \left[(2a_{11}, \; a_{12}+a_{21}) \begin{bmatrix} x \\ y \end{bmatrix} \quad [(a_{12}+a_{21}), \; 2a_{22}] \begin{bmatrix} x \\ y \end{bmatrix} \right]$$

1.10 Back to least square regression

Minimizing any quadratic function of n-variables.

$$= \left[[x \cdot y] \begin{bmatrix} 2a_{11} \\ a_{12}+a_{21} \end{bmatrix}, \; [x \quad y] \begin{bmatrix} a_{12}+a_{21} \\ 2a_{22} \end{bmatrix} \right]$$

$$= [x \quad y] \begin{bmatrix} 2a_{11} & a_{12}+a_{21} \\ a_{12}+a_{21} & 2a_{22} \end{bmatrix}$$

$$= [x \quad y] \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \right)$$

$$\frac{\partial}{\partial \underline{x}} \underline{x}^T A \underline{x} = \underline{x}^T (A + A^T)$$

if $A = A^T$ $\quad \frac{\partial}{\partial \underline{x}} \underline{x}^T A \underline{x} = 2\underline{x}^T A$

$$\frac{\partial}{\partial \underline{x}} \underline{b}^T \underline{x} = b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

$$\frac{\partial}{\partial \underline{x}} \underline{b}^T \underline{x} = \left[\frac{\partial}{\partial x_1} \underline{b}^T \underline{x} \quad \dots \quad \frac{\partial}{\partial x_n} \underline{b}^T \underline{x} \right]$$

$$= \left[b_1 , b_2 \dots b_n \right] = b^T$$

$$\boxed{\frac{\partial}{\partial \underline{x}} \underline{b}^T \underline{x} = b^T}$$

$$\boxed{\frac{\partial}{\partial \underline{x}} \underline{x}^T A \underline{x} = 2\underline{x}^T A}$$

$$\frac{\partial}{\partial x} ax = a$$

$$\frac{\partial}{\partial x} ax^2 = 2ax$$

$$= y^T y + \underline{m}^T X^T X \underline{m} - 2\underline{y}^T X \underline{m}$$

$$\underset{\underline{m}}{\arg\min}\, e(\underline{m}) = \underline{m}^T A \underline{m} + \underline{b}^T \underline{m} + d$$

$A = X^T X$

$b^T = -2y^T X$

$d = y^T y$

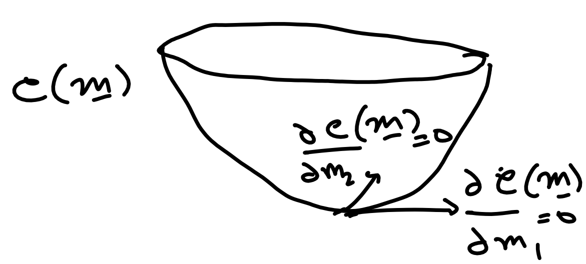
$$\frac{\partial e(\underline{m})}{\partial \underline{m}} = 0^T$$

$$\underbrace{\frac{\partial}{\partial \underline{m}} \underline{m}^T A \underline{m}} + \underbrace{\frac{\partial}{\partial \underline{m}} \underline{b}^T \underline{m}} + \frac{\partial}{\partial \underline{m}} d = \underline{0}^T$$

$$2\underline{m}^T A + \underline{b}^T + 0 = 0^T$$

$$2A^T \underline{m} + \underline{b} = 0$$

if A^T is invertible then $\underline{m} = -\frac{(A^T)^{-1}\underline{b}}{2}$



$$A^T = A = \underbrace{X^T X}$$

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{bmatrix} \in \mathbb{R}^{n \times 2}$$

$$\underline{b}^T = -2 y^T X$$

$$\underline{b} = -2 X^T \underline{y}$$

$$X^T \in \mathbb{R}^{2 \times n}$$

$$\underset{2 \times n}{X^T} \quad \underset{n \times 2}{X}$$

$$\underset{m \times n}{A} \quad \underset{n \times p}{B}$$

$$AB \quad m \times p$$

$$\boxed{\underline{m} = \underbrace{(X^T X)^{-1} X^T} \underline{y}}$$

$$X^T X \in 2 \times 2$$

$$(X^T X)^T = X^T (X^T)^T = X^T X$$

$$(X^T X)^{-1} X^T = X^\dagger = \text{Pseudo inverse of } X$$

1.11 Code in numpy