Fitting Models to Data and Matrix Multiplication

Objectives

- introduce notation for matrices
- review matrix multiplication
- data modeling using matrix multiplication
- introduce block matrix multiplication

A matrix is a collection of values arranged in rows and columns

upper case underscore

$$\underline{A} = \begin{bmatrix} 3 & 6 \\ 2 & 5 \\ 1 & 4 \end{bmatrix}$$

3 rows

2 columns

$\underline{A}$ is a 3 by 2 matrix

☆ $[\underline{A}]_{ij}$: element in row i, column j

$$[\underline{A}]_{2,1} = 2$$

Let $\underline{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{Li} \end{bmatrix}, i = 1, 2, \cdots M$

$L \times 1$

columns

$$\underline{B} = \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_M \end{bmatrix}$$

$L \times M$

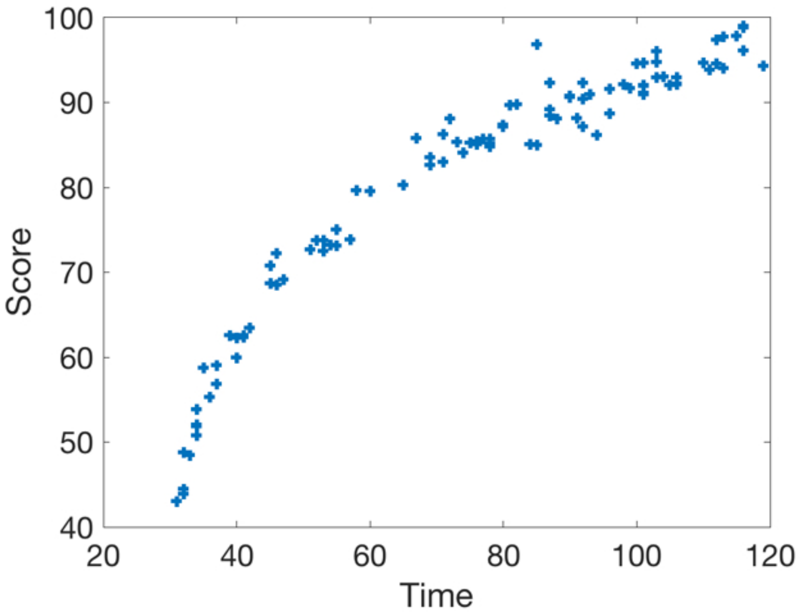
$$\underline{C} = \begin{bmatrix} \underline{a}_1^T \\ \underline{a}_2^T \\ \vdots \\ \underline{a}_M^T \end{bmatrix}$$

rows

$M \times L$

$$\underline{C} = \underline{B}^T$$

Matrices are used to model data



prediction

$$\hat{s} = w_1 + w_2 t + w_3 t^2$$

$$= \begin{bmatrix} 1 & t & t^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \underline{t}^T \underline{w}$$

inner product

Each single data pair

Data

$(s_i, t_i) \quad i = 1, 2, \cdots N$

Find $\underline{w}$

$$\underline{t}_i^T = \begin{bmatrix} 1 & t_i & t_i^2 \end{bmatrix}$$

"feature"

s_i "label" $\longrightarrow$ Combine data $\longrightarrow$

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{bmatrix} = \begin{bmatrix} \underline{t}_1^T \\ \underline{t}_2^T \\ \vdots \\ \underline{t}_N^T \end{bmatrix} \underline{w}$$

or $\underset{N \times 1}{\underline{s}} = \underset{N \times 3}{T} \, \underset{3 \times 1}{\underline{w}}$

Matrix Multiplication

$\underline{A} : [\underline{A}]_{ij}$ $\quad i = 1, 2, \cdots, N$, $\quad j = 1, 2, \cdots, M$

$N \times M$

$\underline{B} : [\underline{B}]_{k\ell}$ $\quad k = 1, 2, \cdots M$, $\quad \ell = 1, 2, \cdots L$

$M \times L$

$$\underline{C} = \underline{A}\,\underline{B} : \quad [\underline{C}]_{mn} = \sum_{j=1}^{M} [\underline{A}]_{mj} [\underline{B}]_{jn}$$

$N \times L$

inner product of m^{th} row of $\underline{A}$ with n^{th} column of $\underline{B}$

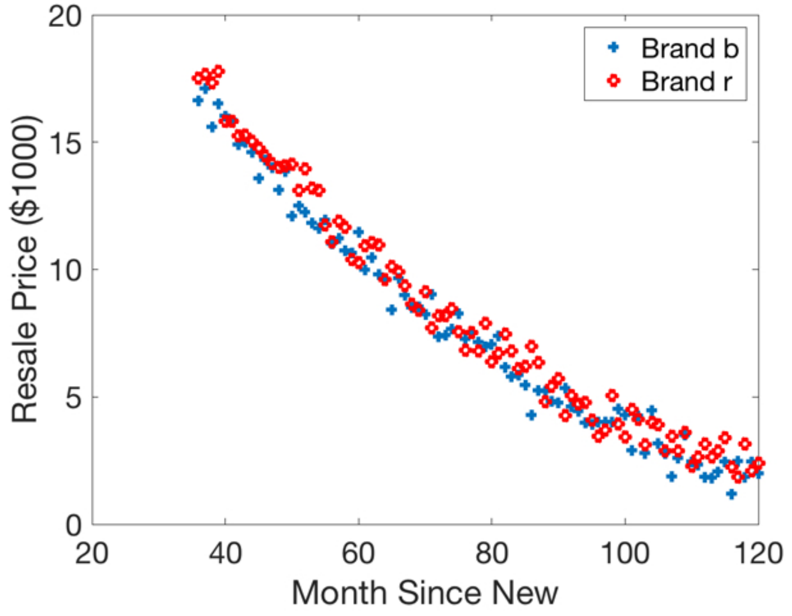
$$\underline{A} = \begin{bmatrix} 3 & 4 \\ 2 & 5 \\ 1 & 6 \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} -2 & 8 \\ 7 & -3 \end{bmatrix}$$

3×2 $\qquad$ 2×2

$$\underline{C} = \begin{bmatrix} 3(-2) + 4 \cdot 7 & 3 \cdot 8 + 4(-3) \\ 2(-2) + 5 \cdot 7 & 2 \cdot 8 + 5(-3) \\ 1(-2) + 6 \cdot 7 & 1 \cdot 8 + 6(-3) \end{bmatrix}$$

3×2

Example: Modeling multiple responses



r: $p_{r_i} = \begin{bmatrix} 1 & t_i & t_i^2 \end{bmatrix} \begin{bmatrix} w_{r_1} \\ w_{r_2} \\ w_{r_3} \end{bmatrix} = \underline{t}_i^T \underline{w}_r$

Features the same

b: $p_{b_i} = \begin{bmatrix} 1 & t_i & t_i^2 \end{bmatrix} \begin{bmatrix} w_{b_1} \\ w_{b_2} \\ w_{b_3} \end{bmatrix} = \underline{t}_i^T \underline{w}_b$

Find $\underline{w}_r, \underline{w}_b$

r: $\begin{bmatrix} p_{r_1} \\ p_{r_2} \\ \vdots \\ p_{r_L} \end{bmatrix} = \begin{bmatrix} \underline{t}_1^T \\ \underline{t}_2^T \\ \vdots \\ \underline{t}_L^T \end{bmatrix} \underline{w}_r$ $\quad \underline{T}$

b: $\begin{bmatrix} p_{b_1} \\ p_{b_2} \\ \vdots \\ p_{b_L} \end{bmatrix} = \begin{bmatrix} \underline{t}_1^T \\ \underline{t}_2^T \\ \vdots \\ \underline{t}_L^T \end{bmatrix} \underline{w}_b$ $\quad \underline{T}$

r+b:

Feature matrix

$$\underbrace{\begin{bmatrix} p_{r_1} & p_{b_1} \\ p_{r_2} & p_{b_2} \\ \vdots & \vdots \\ p_{r_L} & p_{b_L} \end{bmatrix}}_{L\times 2} = \underbrace{\underline{T}}_{L\times 3} \underbrace{\begin{bmatrix} \underline{w}_r & \vdots & \underline{w}_b \end{bmatrix}}_{3\times 2}$$

Multiplication rules extend to block matrices

Previous example: $\left[\underline{P}_r \vdots \underline{P}_b\right] = \underline{I}\left[\underline{W}_r \vdots \underline{W}_b\right]$ → partitioning into blocks

Generalizing $\underline{A} = \left[\begin{array}{c:c} \underline{A}_{11} & \underline{A}_{12} \\ \hdashline \underline{A}_{21} & \underline{A}_{22} \end{array}\right]$, $\underline{B} = \begin{bmatrix} \underline{B}_{11} & \underline{B}_{12} \\ \underline{B}_{21} & \underline{B}_{22} \end{bmatrix}$

$$\underline{C} = \underline{A}\,\underline{B} = \left[\begin{array}{c:c} \underline{A}_{11} & \underline{A}_{12} \\ \hdashline \underline{A}_{21} & \underline{A}_{22} \end{array}\right]\left[\begin{array}{c:c} \underline{B}_{11} & \underline{B}_{12} \\ \hdashline \underline{B}_{21} & \underline{B}_{22} \end{array}\right] = \left[\begin{array}{c:c} \underline{A}_{11}\underline{B}_{11} + \underline{A}_{12}\underline{B}_{21} & \underline{A}_{11}\underline{B}_{12} + \underline{A}_{12}\underline{B}_{22} \\ \hdashline \underline{A}_{21}\underline{B}_{11} + \underline{A}_{22}\underline{B}_{21} & \underline{A}_{21}\underline{B}_{12} + \underline{A}_{22}\underline{B}_{22} \end{array}\right]$$

☆ All relevant submatrices must be conformable → (Matrix multiplication exists)

$\underline{A}_{11}\underline{B}_{11}$, $\underline{A}_{12}\underline{B}_{21}$, $\underline{A}_{21}\underline{B}_{11}$, ... must be defined

[☆ You can partition matrix into blocks of other matrices/vectors]