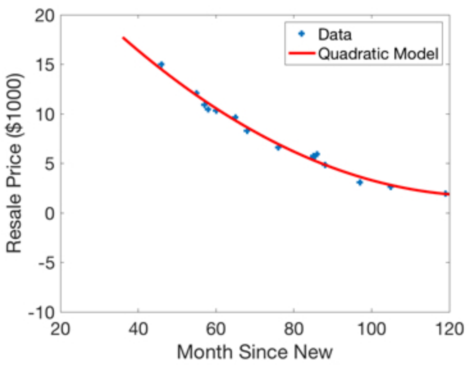
Linear Independence and Rank in Learning

Objectives

- Review the role of systems of linear equations in machine learning
- Define linear independence
- Define the rank of a matrix

Learning classifiers and data models requires solving systems of linear equations

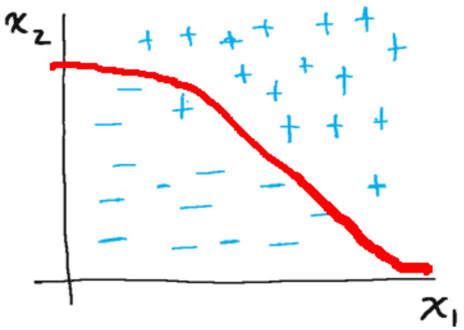


$$\begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \vdots \\ \hat{p}_{20} \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ & \vdots & \\ 1 & t_{20} & t_{20}^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

Model fitting

$$\underline{A}\,\underline{w} = \underline{d}$$

$(N \times M)\,(M \times 1)\;(N \times 1)$



$$\begin{bmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_N \end{bmatrix} = \begin{bmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \vdots \\ \underline{x}_N^T \end{bmatrix} \underline{w}$$

Label

Classifier design

Find set of ws

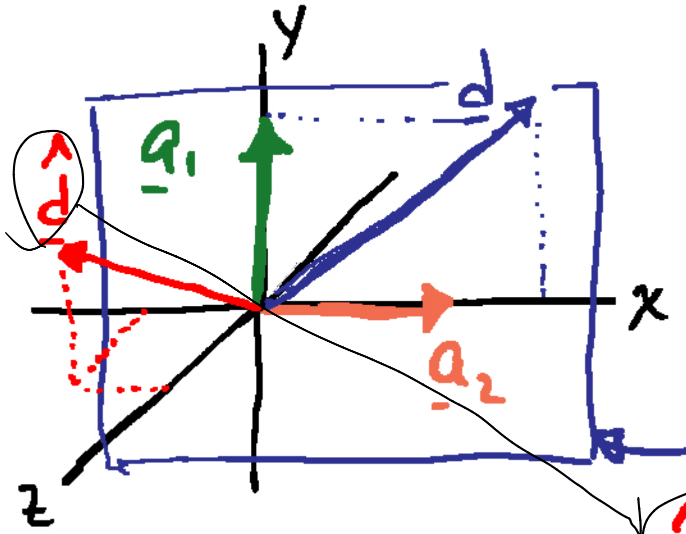
Important: Can we solve $\underline{A}\,\underline{w} = \underline{d}$?

Let $\underline{A} = \left[\underline{a}_1\ \underline{a}_2\ \cdots\ \underline{a}_M\right]$ so $\underline{A}\underline{w} = \underline{d} \Rightarrow \underline{d} = \underbrace{\sum_{i=1}^{M} \underline{a}_i w_i}_{\text{weighted sum of } \underline{a}_i}$

$N \times 1$

Example: $\underline{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$

$\begin{array}{cc} \underline{a}_1 & \underline{a}_2 \end{array}$



$\underline{d} = \underline{a}_1 w_1 + \underline{a}_2 w_2$

set of all possible $\underline{d}$

$\hat{\underline{d}} \neq \underline{a}_1 w_1 + \underline{a}_2 w_2$

(Not in the triangle area: There is no z-axis component)

For a solution $\underline{d} - \sum_{i=1}^{M} \underline{a}_i w_i = \underline{0}$

☆ **Linear independence:** A set of M vectors $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_M \in \mathbb{R}^N$ is linearly independent iff $\sum_{i=1}^{M} \underline{v}_i \alpha_i = 0 \Leftrightarrow \alpha_i = 0, i = 1, 2, \cdots, M$

(Their weights are 0 → the sum is 0)

otherwise "linearly dependent"

☆ **Rank of a matrix:** max number of linearly independent columns (or rows) it has

☆ Note: row rank = column rank

Examples:

$$\underline{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$
$$\quad \underline{a}_1 \quad \underline{a}_2 \quad \underline{a}_3$$

$\underline{a}_3 = \underline{a}_2 - \underline{a}_1$

so $\alpha_1 \underline{a}_1 + \alpha_2 \underline{a}_2 + \alpha_3 \underline{a}_3 = \underline{0}$

for $\alpha_1 = 1, \alpha_2 = -1, \alpha_3 = 1 \neq 0$

$\underline{a}_1, \underline{a}_2, \underline{a}_3$ are linearly dependent!

$\alpha_1 \underline{a}_1 + \alpha_2 \underline{a}_2 = 0$,

$\Updownarrow$

$\alpha_1 = \alpha_2 = 0$

$\underline{a}_1, \underline{a}_2$ lin. indep.

$\alpha_1 \underline{a}_1 + \alpha_3 \underline{a}_3 = 0$

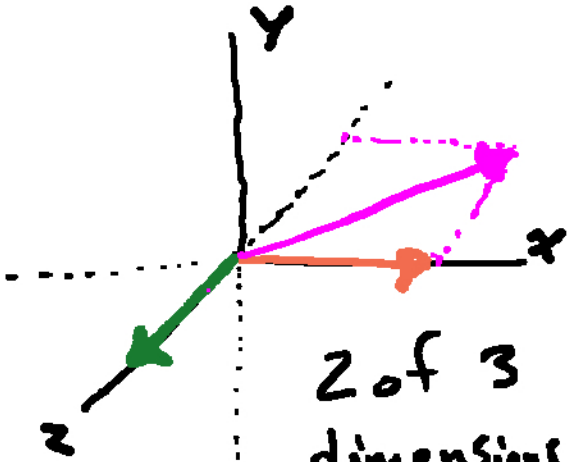
$\Updownarrow$

$\alpha_1 = \alpha_3 = 0$

$\underline{a}_1, \underline{a}_3$ lin indep

same for $\underline{a}_2, \underline{a}_3$

$\mathrm{rank}(\underline{A}) = 2$



2 of 3 dimensions

☆ Rank 0 $\rightarrow$ The matrix has no linearly independent columns.
(Only happens with the zero matrix)

Rank 1 $\rightarrow$ At least one non-zero column, and all other columns are scalar multiples of that one.

(☆ If A is $n \times p$, $\text{rank}(A) = p$

then $\text{rank}(A^TA) = p$, A^TA is invertible)

☆ $\because$ rank of columns = rank of rows of a matrix

$\therefore \text{rank}(A) = \text{rank}(A^T)$

$\because$ B is $n \times n$ / invertible matrix

$\therefore B^{-1}$ is also $n \times n$ / invertible

$\downarrow$

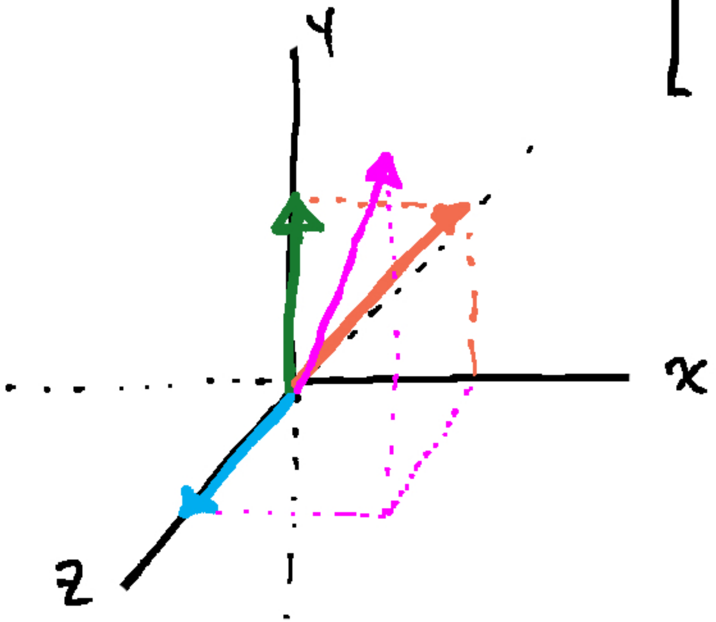
$\text{rank}(B^{-1}) = \text{rank}(B) = n$

Example: $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$
$\underline{a}_1 \quad \underline{a}_2 \quad \underline{a}_3 \quad \underline{a}_4$

$\sum_{i=1}^{4} \alpha_i \underline{a}_i = \underline{0} \not\Leftrightarrow \alpha_i = 0$ 5

$\underline{a}_1 + \underline{a}_2 - \underline{a}_3 + \underline{a}_4 = \underline{0}$

$\underline{a}_1, \underline{a}_2, \underline{a}_3, \underline{a}_4$ lin. dep.



$\alpha_1 \underline{a}_1 + \alpha_2 \underline{a}_2 + \alpha_3 \underline{a}_3 = 0 \Leftrightarrow \alpha_i = 0$

any combination of 3 are lin. indep.

$\text{Rank}(\underline{A}) = 3$ ☆ (Because a_1/a_2 has no z-axis component)

any 3 vectors describe all 3 dimensions

Back to $\underline{A}\underline{w} = \underline{d}$ solution: $\underline{A}\underline{w} - \underline{d} = \underline{0}$ or

$\sum_{i=1}^{m} \underline{a}_i w_i + (-1)\underline{d} = \underline{0} \Rightarrow \underline{a}_1, \dots, \underline{a}_m, \underline{d}$ are lin. dep.

$Aw = d$

$\hookrightarrow w_1a_1 + w_2a_2 + \cdots + w_m a_m = d$

If this has solution, then it indicates d can be represented by linear combinations of A's columns

$\therefore$ d is redundant and not providing any new direction

$\therefore$ $w_1a_1 + w_2a_2 + \cdots + w_m a_m - d = 0$

$\therefore$ set of vectors now is $\{a_1, a_2 \cdots a_m, d\}$

set of weights now is $\{w_1, w_2 \cdots w_m, -1\}$

$\because -1 \neq 0$

$\therefore$ when $Aw = d$ has solutions, $\{a_1, a_2 \cdots a_m, d\}$ must be linearly dependent

Summary

$$\underline{A}\,\underline{w} = \underline{d} \qquad \underline{A} = [\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_M]$$

1) If $\underline{d}$ is a linear combination of $\underline{a}_1, \underline{a}_2, \cdots \underline{a}_M$ there is a solution

$$\text{rank}(\underline{A}) = \text{rank}([\underline{A} \vdots \underline{d}])$$

2) If $\underline{a}_1, \underline{a}_2, \cdots \underline{a}_M, \underline{d}$ are linearly independent, there is no solution

$$\text{rank}(\underline{A}) < \text{rank}([\underline{A} \vdots \underline{d}])$$

3) Unique solution ?