MAL 7: L.A in ML

Linear equations:

Ex:

Ex: (Plane equation)

A system of Linear Equations

A collection of one or more linear equations involving the same variables:

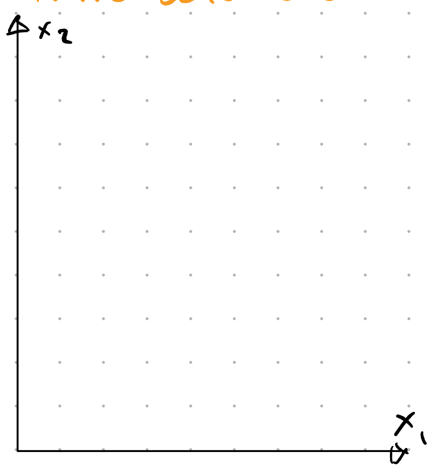
Ex:

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 3 \\ 7x_2 - 4x_3 &= 10 \\ x_3 &= 1 \end{aligned}$$

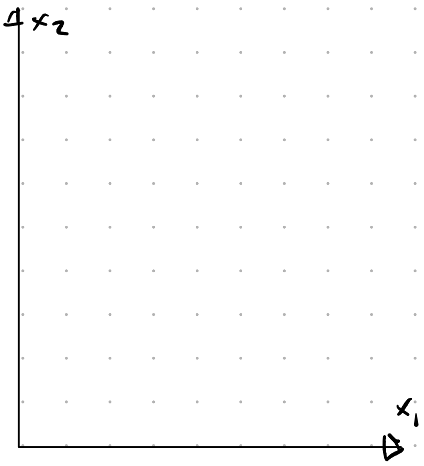
Solution set:

A solution of a linear system is a list of numbers $s_1, s_2, s_3 \dots$ that satisfies the system, i.e. makes the system "true".

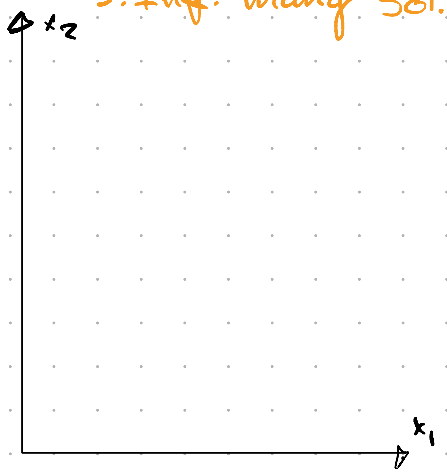
1. No solution



2. One solution



3. Inf. many sol.



The Matrix:

Consider the system:

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 3 \\ 7x_2 - 4x_3 &= 10 \\ x_3 &= 1 \end{aligned}$$

We can "code" this system into two types of matrices:

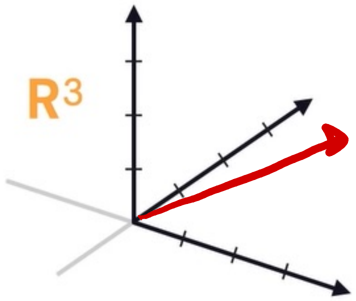
Coefficient Matrix

Just think of a Matrix as a 2-D array.

The Vector:

- A vector is a 1-D array of n numbers
- Vector Space: The totality of all vectors with n entries is an n-dimensional vector space

$$\vec{v} = \nearrow = \begin{bmatrix} -3 \\ 0.7 \\ 2 \end{bmatrix}$$

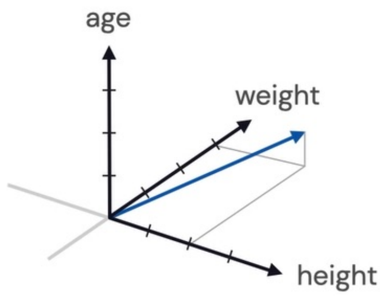


"3-dimensional space" consists of all vectors with 3 entries:

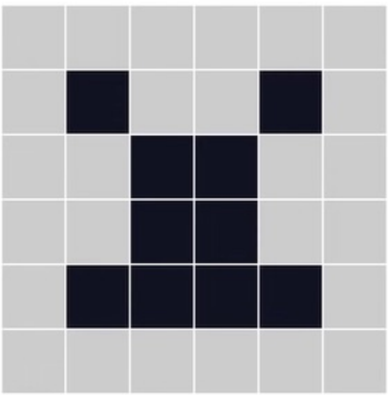
$$\begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

- A feature vector is a vector whose entries represent the features of some object:

$$P = \begin{array}{c} \begin{bmatrix} 167 \\ 60 \\ 23 \end{bmatrix} \end{array} \begin{array}{l} \text{height} \\ \text{weight} \\ \text{age} \end{array}$$



Images:

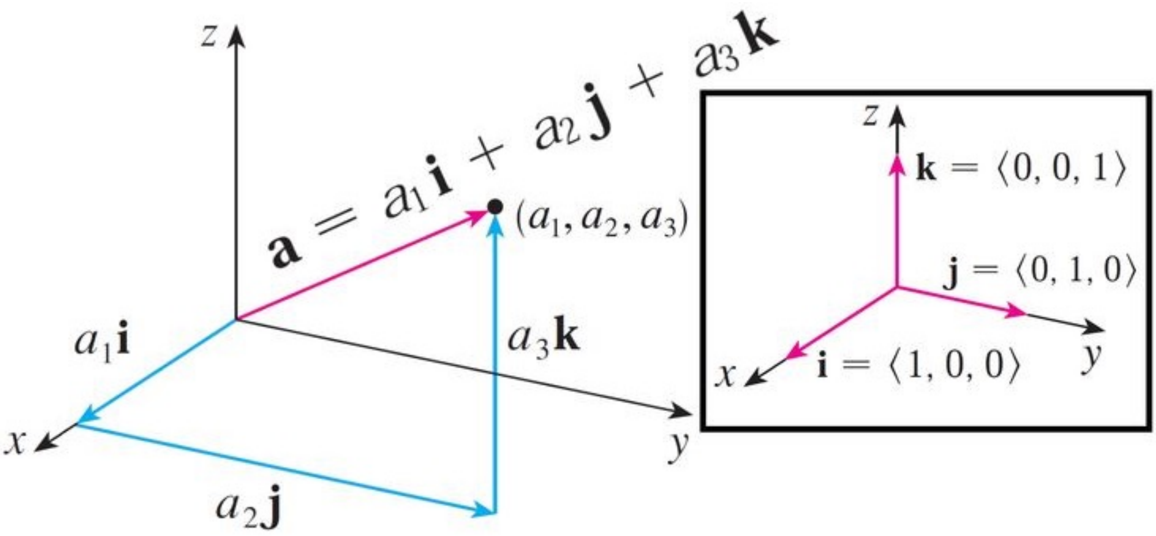


$$\rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

One hot-encoding:

$$\text{apple} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{cat} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{house} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{tiger} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

In math, these are called standard basis vectors:



Any vector can be stated as a linear combination of some basis vector.

linear Combinations:

Given a set of vectors $\bar{v}_1, \bar{v}_2 \dots \bar{v}_p \in \mathbb{R}^n$ and scalars $c_1, c_2 \dots c_p \in \mathbb{R}$, the vector $\bar{y}$ given by:

is called a linear combination of $\bar{v}_1, \bar{v}_2 \dots \bar{v}_p$ with weights $c_1, c_2 \dots c_p$.

The Matrix:

Denoted with capital letters, e.g. A, X

- If you "read" the matrix horisontally you see equation.
- If you "read" the matrix vertically, you see vectors

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

	X_1 sepal length	X_2 sepal width	X_3 petal length	X_4 petal width	X_5 class
x_1	5.9	3.0	4.2	1.5	Iris-versicolor
x_2	6.9	3.1	4.9	1.5	Iris-versicolor
x_3	6.6	2.9	4.6	1.3	Iris-versicolor
x_4	4.6	3.2	1.4	0.2	Iris-setosa
x_5	6.0	2.2	4.0	1.0	Iris-versicolor
x_6	4.7	3.2	1.3	0.2	Iris-setosa
x_7	6.5	3.0	5.8	2.2	Iris-virginica
x_8	5.8	2.7	5.1	1.9	Iris-virginica
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{149}	7.7	3.8	6.7	2.2	Iris-virginica
x_{150}	5.1	3.4	1.5	0.2	Iris-setosa

The number of rows tell you the size of of the vectors:

Matrix Properties

The ith row : $A_{i,:}$

The jth col : $A_{:,j}$

transpose A^T : $(A^T)_{i,j} = A_{j,i}$

Vectors are matrices with one column.

Multiplication:

$$C = A \cdot B \Rightarrow C_{ij} = \sum_k A_{ik} \cdot B_{kj}$$

Example:

Let $A = 5 \times 2$:

$$A^T A =$$

$$A A^T =$$

Matrix and vectors:

$$A\bar{x} = \bar{b}, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad \bar{x} \in \mathbb{R}^n$$

$$A_{1,:} \cdot \bar{x} = b_1, \quad A_{2,:}\, \bar{x} = b_2 \quad \dots \quad A_n\, \bar{x} = b_m$$

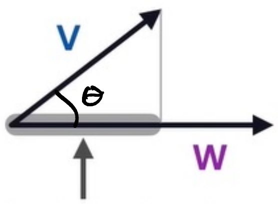
$$A\, \bar{x} = \bar{b} \Rightarrow A^{-1} \cdot \bar{x} = A^{-1}\bar{b} \Rightarrow \bar{x} = A^{-1}\bar{b}$$

$$x = y \cdot a^{-1}$$

Dot-Product (Inner Product):

$$\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 2 \\ -1 \end{bmatrix} =$$

The dot product says something about how similar to vectors are



Length of shadow is **V•W**
(if length of **W** is **1**)

$$\bar{v} \cdot \bar{w} = \bar{v}^T \times \bar{w} = \|\bar{v}\|_2 \cdot \|\bar{w}_2\| \cdot \cos\theta$$

Note: if X is your data XX^T is called the Covariance matrix since it will contain all the dot products.

Matrices of Functions:

It is useful to think of a matrix as a function that "acts" on a vector:

$f(x)$ ↳ Turn $\bar{v}$ into $\bar{v}'$: stretch, rotate, scale

$T(\bar{x}) = A\bar{x}$ ← transformation: $T: \mathbb{R}^n \to \mathbb{R}^m$

Example:

$$\bar{x} = \begin{bmatrix} 5 \\ 6 \\ -2 \end{bmatrix}, \quad A = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 4 & 5 \end{bmatrix}$$

$$T(\bar{x}) = \begin{bmatrix} 2 & 0 & -3 \\ -1 & 4 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ -2 \end{bmatrix} =$$

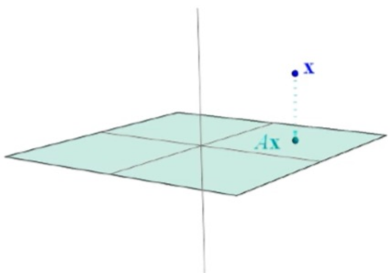
So we transformed a 3-D object to a 2-D object.

Special Transformations:

- Consider $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and let $T(\boldsymbol{x}) = A\boldsymbol{x}$.

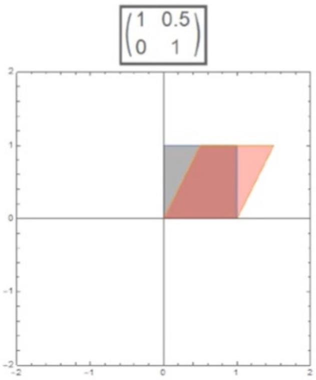
Projection

- Now $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$ "projects" $\mathbb{R}^3$ onto the xy-plane



Consider $A = \begin{bmatrix}1 & c \\ 0 & 1\end{bmatrix}$, where c is a real number

Shear



Rotation

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta)\end{bmatrix}$$

=

To rotate counterclockwise by:

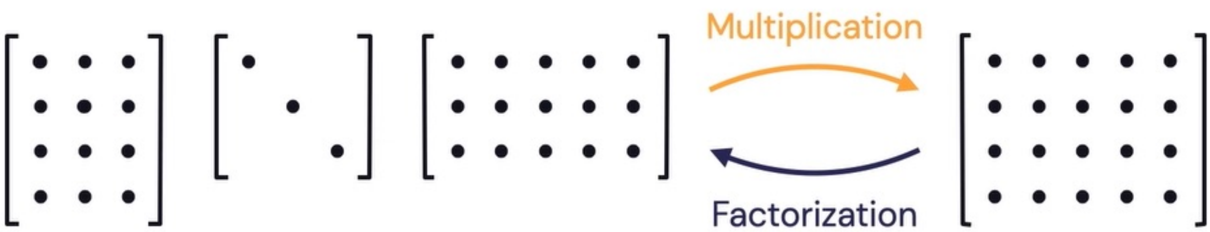
	90 degrees	180 degrees	270 degrees
Multiply by:	$\begin{bmatrix}0 & -1 \\ 1 & 0\end{bmatrix}$	$\begin{bmatrix}-1 & 0 \\ 0 & -1\end{bmatrix}$	$\begin{bmatrix}0 & 1 \\ -1 & 0\end{bmatrix}$

Covariance

Covariance Matrix

$$\frac{AA^T}{n-1} = \begin{bmatrix}\mathrm{Var}(x_1) & \cdots & \mathrm{Cov}(x_n, x_1) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(x_n, x_1) & \cdots & \mathrm{Var}(x_n)\end{bmatrix} \to \text{Corr. mat.}$$

Matrix Decomposition:



Eigendecomposition

$A\bar{x}$

An eigenvector is a special vector s.t.

$$A\bar{v} = \lambda\bar{v}$$

Where λ is called the eigenvalue.
scale factor

If A has n eigenvectors $\{\bar{v}_1, \bar{v}_2, \bar{v}_3 \dots \bar{v}_n\}$
with corresponding eigenvectors
$\lambda = [\lambda_1, \lambda_2, \lambda_3 \dots \lambda_n]$
Then then the eigendecomposition of A
is

$$A = V \cdot D \cdot V^{-1}$$

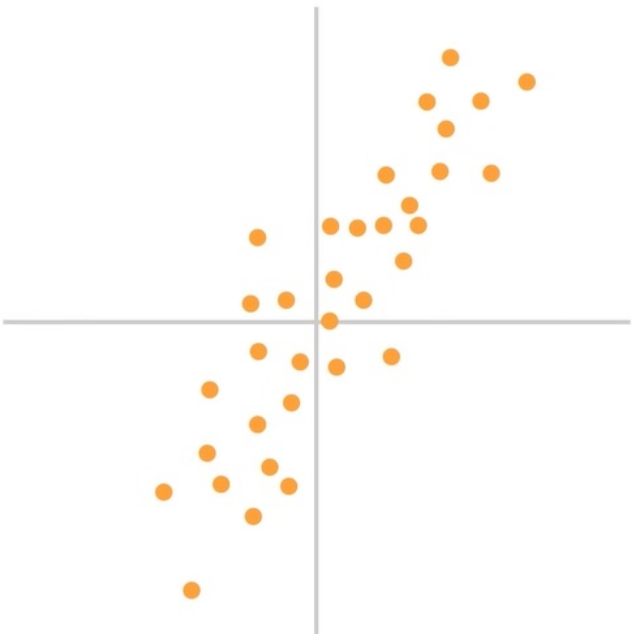
$$V = \begin{bmatrix} | & | & | & & | \\ \bar{v}_1 & \bar{v}_2 & \bar{v}_3 & \dots & \bar{v}_n \\ | & | & | & & | \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Note $V \cdot V^{-1} = I$

Example:

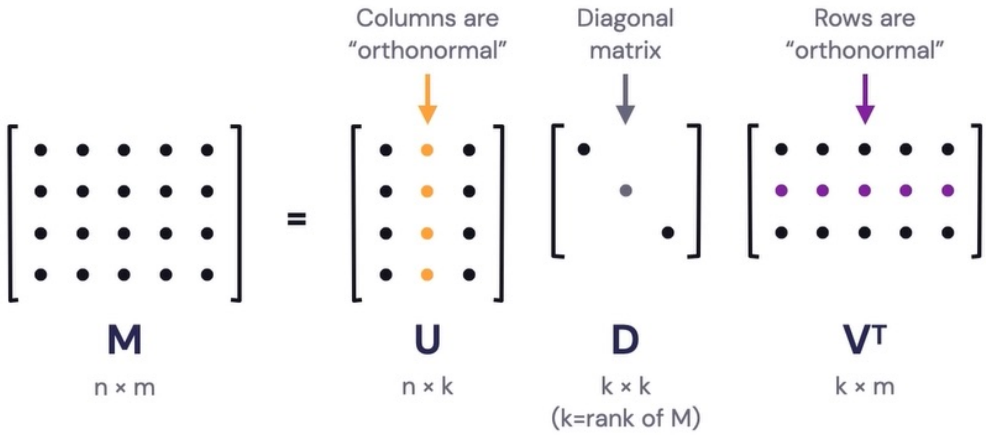
$$\bar{v}_1 = \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}, \quad \bar{v}_2 = \begin{bmatrix} -1.4 \\ 1 \end{bmatrix}$$

$$\lambda_1 = 57, \quad \lambda_2 = 3$$



Singular Value Decomposition:

singular value = $\sqrt{\text{eigenvalue}}$



Used in:

- Linear Regression
- PCA
- Image Compression
- Recommendation Systems
- Classification of handwritten digits
- Restoring images