UNIVERSIDAD PANAMERICANA

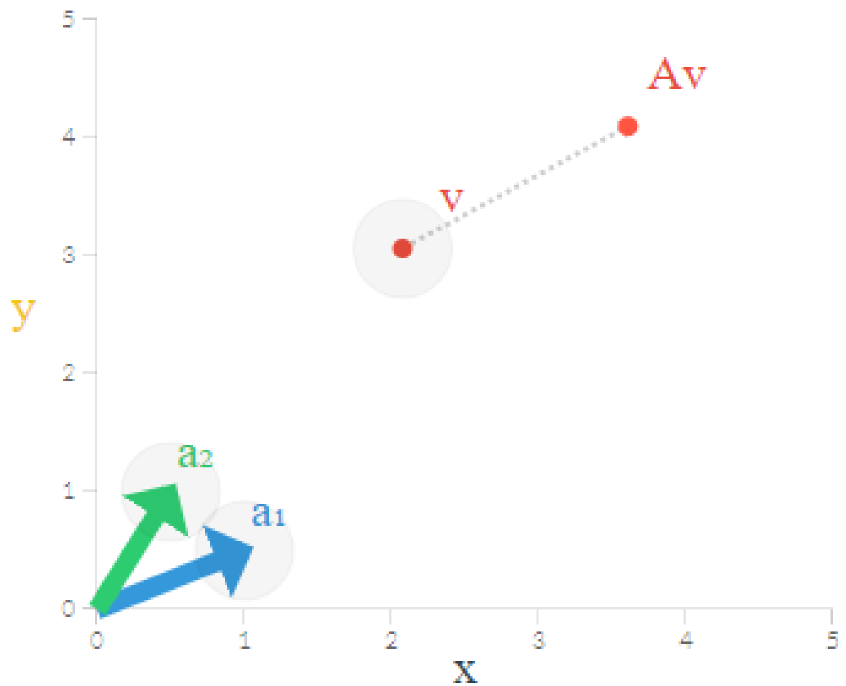
Machine Learning

(https://leonpalafox.github.io/mlclase/)

Leon F. Palafox PhD

Eigenvectores

V es un vector (mostrado como un punto) y A una matriz con columnas a1 y a2 (como flechas). Si multiplicamos v por A, entonces A convierte v en un nuevo vector Av.



$$A = \begin{bmatrix} a_{1,x} & a_{2,x} \\ a_{1,y} & a_{2,y} \end{bmatrix} = \begin{bmatrix} 1.00 & 0.50 \\ 0.50 & 1.00 \end{bmatrix}$$

$$v = \begin{bmatrix} 2.08 \\ 3.06 \end{bmatrix}$$

$$Av = \begin{bmatrix} 3.61 \\ 4.10 \end{bmatrix}$$

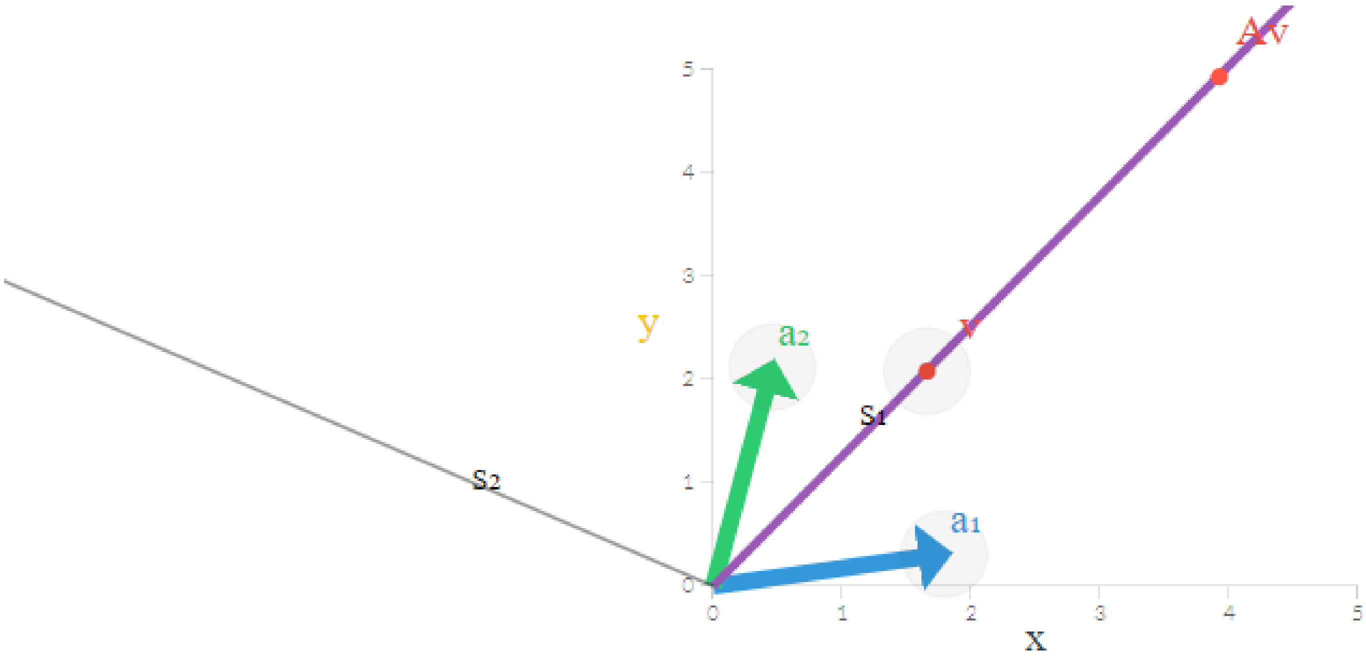
Eigenvectores

- Si pudieses dibujar una línea a través de los 3 puntos (0,0), v y Av, entonces Av es simplemente v multiplicado por un valor λ; que se expresa como, Av=λv. En este caso, llamamos λ un eigenvalor y v un eigenvector. For example, here (1,2) is an eigvector and 5 an eigenvalue.

$$Av = \begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \lambda v.$$

Eigenvectores

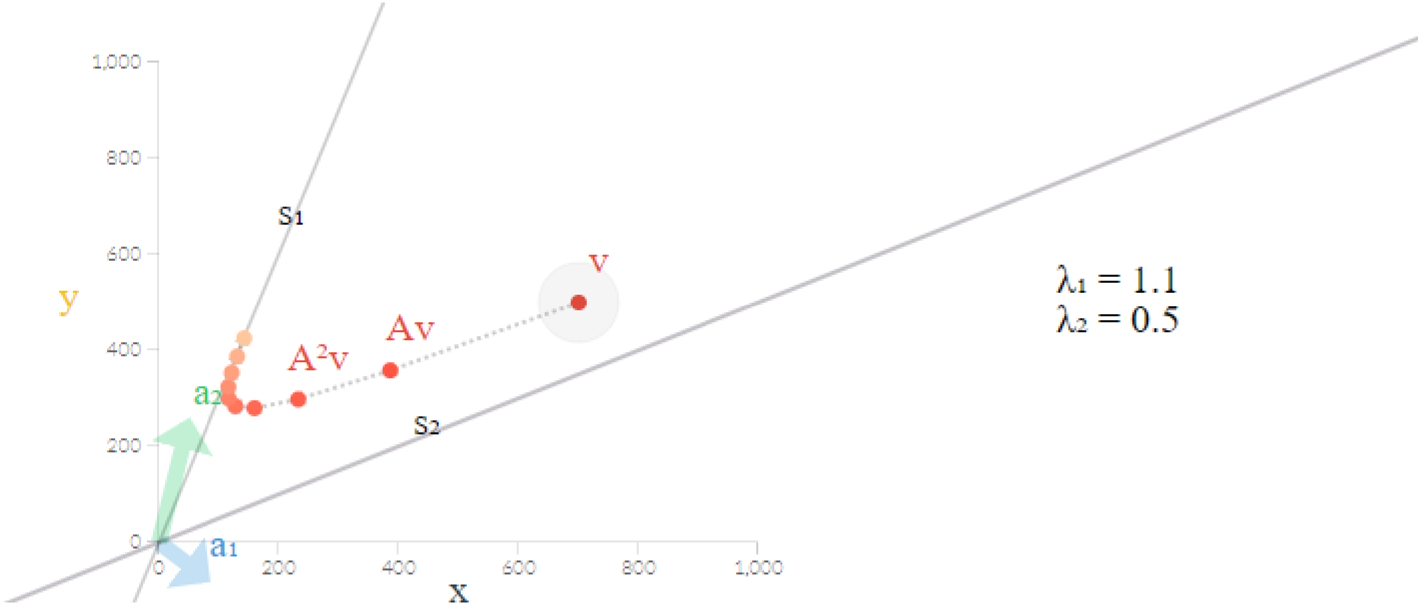
Hay multiples vectores V que pueden cumplir las características (eigenespacios)



$\lambda_1 = 2.37$
$\lambda_2 = 1.55$

Eigenvalores

Si seguimos multiplicando v por A, obtienes una secuencia v,Av,A*A*v, etc. Los eigenespacios son atractores y los eigenvalores te dicen si termina en (0,0) o no. Los eigenvalores nos hablan de la naturaleza de evolución de un sistema.



Eigenvalores

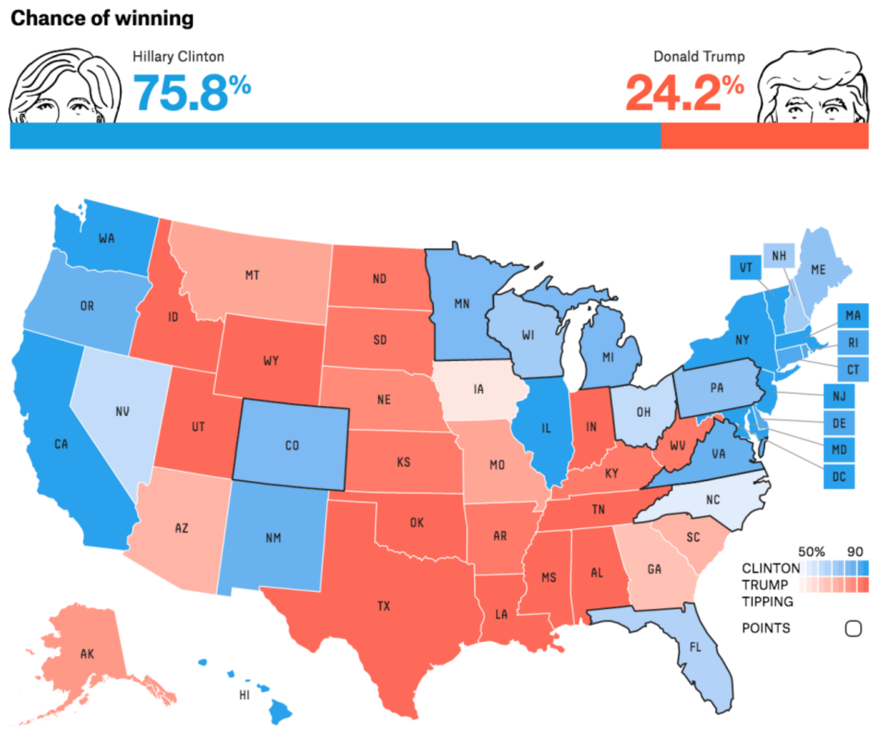
- Ejemplos:
  - http://setosa.io/ev/eigenvectors-and-eigenvalues/
  - https://github.com/leonpalafox/mlclase/blob/master/Chapter2Fundamentals/MatrixAlgebra.ipynb

Probabilidad

Probabilidad de que suceda un evento

P = 1/12

- La palabra probabilidad tiene diferentes connotaciones



En el caso de México

- http://oraculus.mx/poll-of-polls/

Distribuciones de Probabilidad

- Que es una probabilidad?
  - Es una medida de la incertidumbre (o certidumbre) de un evento.
- En Machine Learning por lo general tratamos con escenarios inciertos.
  - El mundo es incierto.
- Tenemos dos tipos de probabilidad
  - Frequentista
  - Bayesiana

Probabilidad Frequentista

- Podemos "crear" tantos escenarios como queramos
  - Bootstrapping
- El ejemplo de los dados
- Partidos de Fútbol (chance)
- Control de Calidad(6 Sigma)
  - 6 desviaciones estandard.
- Por lo general se usa mal.

Probabilidad Bayesiana

- No podemos crear tantos escenarios como queramos
  - Accidentes de avión
  - Elecciones prescidenciales
  - Diagnósticos médicos
- Aqui, la probabilidad representa un grado de creencia en un evento.
- No se usa lo suficiente.

Comparten las mismas reglas

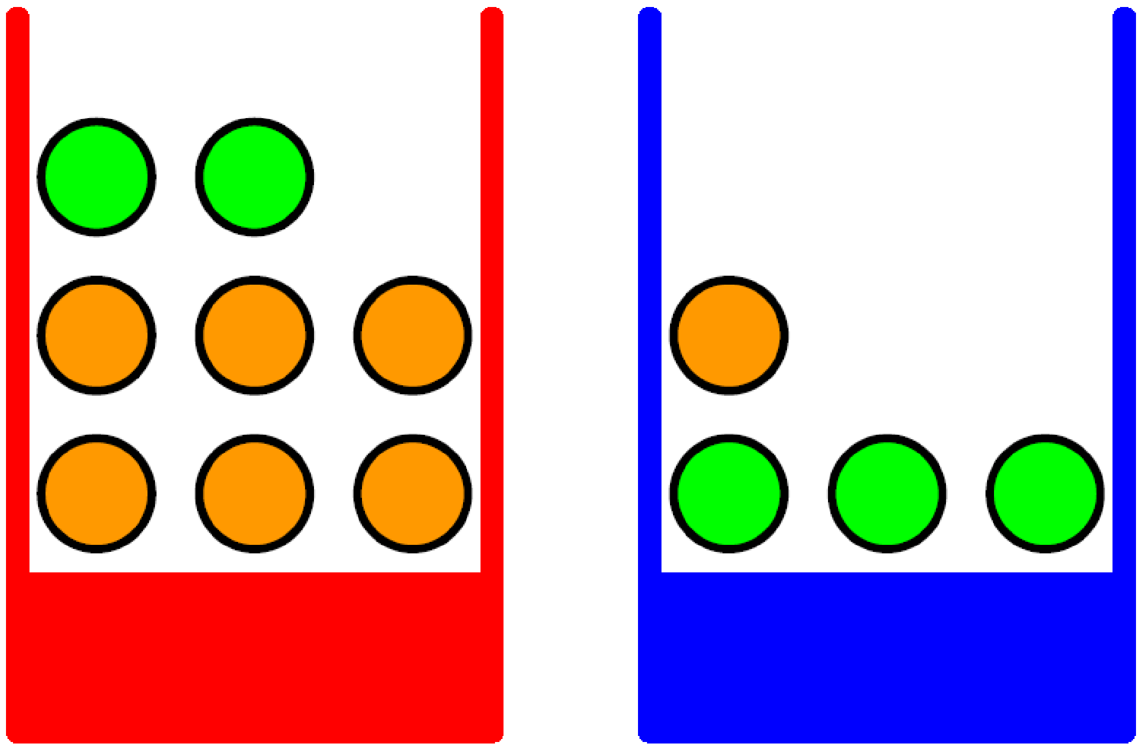
- Ambos, Bayesianos y Frequentistas comparten los mismos principios
  - Math
  - Leyes
  - Teoremas.
- Es facil mezclarlos
  - Cuidado con los P-values!!!!!!!

Muestreo aleatorio

- Equipos Fútbol
  - Messi, Agüero, Higüain
  - Ronaldo, Bale, Navas
  - Messi, Suárez, Paulinho
- Películas
  - Clooney, Damon, Robo
  - Damon, Mars, NASA
  - Damon, Robin Williams, MIT

- Cada equipo/pelicula es una distribución probabilística.

Fruit examples (Bishop's)



Regresion Lineal

- Nuestra Primera Herramienta
- Es una de las herramientas más fáciles de automatizar

Regresión Lineal

- Imaginemos que quieren vender su auto:
  - Cuanto pides por el auto:
    - Km
    - Año
    - Color
    - Opciones
    - Condición

Hay sitios como Edmunds.com

Model	Year	Brand	Price	Options	Condition	Mileage
Corvette	1961	Chevrolette	100K	Standard	As New	100,000
Corvette	1961	Chevrolette	10K	Standard	Rust	100,000
Corvette	1961	Chevrolette	120K	Standard	Used	20,000

Regresión Lineal

Precio = Año + Opciones + Condición + Kms

¿Qué está mal con esta ecuación?

¿Cómo llamamos a las variables Año, Opciones, etc…?

- Como la modificaríamos?

Price = A*Price + B*Options + C*Condition+ D*Mileage

- Que es A, B, C y D?

Que pasó con los features

- Por que nuestra primer ecuación no funcionó?

- Que pasaría si incrementaramos el número de características?

Regresión Lineal

- Estamos tratando de predecir variables continuas
- Tenemos características (features) que en teoría son independientes las unas de las otras.
- Queremos encontrar el mejor conjunto de pesos para resolver el problema.

Aplicaciones para la regression Lineal

- Precios de Casas
- Presupuesto para una película
- Efecto de un tratamiento
- Inclinación política (1-100)
- Número de Likes
- Cantidad de ventas (para un producto)

Formalmente

$$y_1 = \beta_1 x_{11} + \cdots + \beta_p x_{1p}$$
$$y_2 = \beta_1 x_{21} + \cdots + \beta_p x_{2p}$$
$$y_3 = \beta_1 x_{31} + \cdots + \beta_p x_{3p}$$

$$y_n = \beta_1 x_{n1} + \cdots + \beta_p x_{np}$$

Forma Matricial

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ x_{21} & \cdots & x_{2p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} \qquad \mathbf{y} = \mathbf{X}\beta$$

Optimization problem

- Optimize:

$$\mathbf{y} = \mathbf{X}\beta$$

- Score:

$$X\beta - y$$

$$(X\beta - y)^2$$

$$\frac{1}{2}(X\beta - y)^2$$

Dos formas de atacar el problema

- Algebra Matricial(next class)

- Gradiente Descendente(next class)