3. Background

- Linear algebra
  - vectors, matrices, inner products, decompositions, norms
- Probability and statistics
  - random variable / vector, sampling from distribution, expectation, variance

Linear algebra

Vectors

All vectors are columns

$$x \in \mathbb{R}^d, \qquad x = \begin{bmatrix} x[1] \\ x[2] \\ \vdots \\ x[d] \end{bmatrix}, \qquad \cancel{x = (x[1], x[2], \ldots, x[d])}$$

x = np.array([1,2,3])

$x = \boxed{\begin{matrix} 1 \\ 2 \\ 3 \end{matrix}}$

x = (1,2,3)

special vectors

- $x = \mathbf{0}$ (zero vector): $x[j] = 0, \quad j = 1, \ldots, d$
- $x = \mathbf{1}$ (ones vector): $x[j] = 1, \quad j = 1, \ldots, d$

Conventions

For a d-vector $x = (x[1], ..., x[d])$ (we also write $x \in \mathbb{R}^d$)

- $x[j]$ refers to the jth component of x
- Two vectors $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^d$ are equal when

$$x[i] = y[i], \ \forall i = 1, ..., d$$

- Inequalities: $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, c scalar

i --> j

$$\begin{aligned} x \geq c \quad &\iff \quad x[j] \geq c, \quad \forall i = 1, ..., d \\ x > c \quad &\iff \quad x[j] > c, \quad \forall i = 1, ..., d \\ x \geq y \quad &\iff \quad x[j] \geq y[j], \quad \forall i = 1, ..., d \\ x > y \quad &\iff \quad x[j] > y[j], \quad \forall i = 1, ..., d \end{aligned}$$

- e.g. if x_i, $i = 1, ..., m$ are samples, then $x_i[k] = k$th element of <u>vector</u> x_i.

Block vectors

- If $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, $z \in \mathbb{R}^p$,

$$a = \begin{bmatrix} x \\ \cdots \\ y \\ \cdots \\ z \end{bmatrix} \in \mathbb{R}^{m+n+p}$$

= (x,y,z)

"stacked" or "block" or "concatenated" vector np.hstack([x,y,z])

- We may also write this as $a = (x, y, z)$.

- If $m = n = p$, can also stack horizontally

$$B = \begin{bmatrix} x & y & z \end{bmatrix} \in \mathbb{R}^{n \times 3}$$

x, y, and z are the first, second, and third columns of B.

Matrices

Another way of collecting numbers

$$A = \begin{bmatrix} A[1,1] & \cdots & A[1,n] \\ \vdots & \ddots & \vdots \\ A[m,1] & \ldots & A[m,n] \end{bmatrix} \in \mathbb{R}^{m \times n}$$

$\in \mathbb{R}^{3 \times 4}$

- $A = 0$ (zero matrix): $A[i,j] = 0$, $i = 1, ..., m$, $j = 1, ..., n$
- Shape of a matrix $A \in \mathbb{R}^{m \times n}$
  - <u>tall</u> if $m > n$, <u>wide</u> if $m < n$, <u>square</u> if $m = n$
- A^T is the <u>transpose</u> of A

$$A^T = \begin{bmatrix} A[1,1] & \cdots & A[m,1] \\ \vdots & \ddots & \vdots \\ A[1,n] & \ldots & A[n,m] \end{bmatrix} \in \mathbb{R}^{n \times m}$$

- Columns of matrices are vectors in $\mathbb{R}^m$.

Conventions

- Two matrices of same size are equal if every element is equal

$$A = B \iff A[i,j] = B[i,j], \quad \forall i = 1, ..., m, \quad j = 1, ..., n.$$

- Inequalities: $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times n}$, c scalar

$$\begin{aligned}
A \geq c \quad &\iff \quad A[i,j] \geq c, \quad \forall i = 1, ..., m, \quad j = 1, ..., n \\
A > c \quad &\iff \quad A[i,j] > c, \quad \forall i = 1, ..., m, \quad j = 1, ..., n \\
A \geq B \quad &\iff \quad A[i,j] \geq B[i,j], \quad \forall i = 1, ..., m, \quad j = 1, ..., n \\
A > B \quad &\iff \quad A[i,j] > B[i,j], \quad \forall i = 1, ..., m, \quad j = 1, ..., n
\end{aligned}$$

- Special matrix I (identity)

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \qquad AI = IA = A$$

Inner products

x[1]y[1] + x[2]y[2] + ... + x[d] y[d]

$$x, y \in \mathbb{R}^d, \qquad x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_d y_d$$

properties

- $(\alpha x)^T y = \alpha (x^T y)$
- $(x + y)^T z = x^T z + y^T z$

inner = x.T * y
outer = x * y.T

- $x^T y = y^T x$
- $x^T x \geq 0$ (sum of squares)
- $x^T x = 0 \iff x = 0$

Examples

- $\mathbf{1}^T x = x[1] + x[2] + ... + x[d]$ (sum entries)
- $x^T x = x[1]^2 + x[2]^2 + ... + x[d]^2$ (sum squares)

Linear cost

- $p = (p[1], ..., p[d])$ vector of prices
- $q = (q[1], ..., q[d])$ vector of quantities

$$\text{total cost} = p^T q = \sum_i p[i]q[i]$$

Matrix inner product

$$A, B \in \mathbb{R}^{m \times n}, \quad \mathbf{tr}(A^T B) = \sum_{i=1}^{m} \sum_{j=1}^{n} A[i,j]B[i,j] = \mathbf{vec}(A)^T \mathbf{vec}(B)$$

where

- for a square matrix $S \in \mathbb{R}^{n \times n}$, $\mathbf{tr}(S) = \sum_{i=1}^{n} S[i,i]$,
- for a matrix with column vectors $a_1, ..., a_n \in \mathbb{R}^m$

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \in \mathbb{R}^{m \times n}, \qquad \mathbf{vec}(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Functions and mappings

- $f : \mathbb{R}^n \to \mathbb{R}$ means f takes n-vectors and returns a real number
- $f : \mathbb{R}^n \to \mathbb{R}$ is a linear function if

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \qquad \forall \alpha, \beta \in \mathbb{R}, \quad x, y \in \mathbb{R}^n$$

- The mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping; e.g.

$$f(x) = Ax, \qquad A \in \mathbb{R}^{m \times n}$$

is a linear mapping.

Norms

Most common norm is "Euclidean norm", i.e. 2-norm:

$$\|x\|_2 = \sqrt{x[1]^2 + \cdots + x[d]^2} = \sqrt{x^T x}$$

Suppose that a, b, and c are vectors. Then

||(a,b,c)||_2 ^2 = (a,b,c)^T (a,b,c) = a^Ta + b^T b + c^T c

$$\|\underbrace{(a, b, c)}_{\text{\textbackslash vector}}\|_2^2 = \|a\|_2^2 + \|b\|_2^2 + \|c\|_2^2 \qquad \text{(Pythagorean identity)}$$

More generally, norms are any function $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ that satisfy

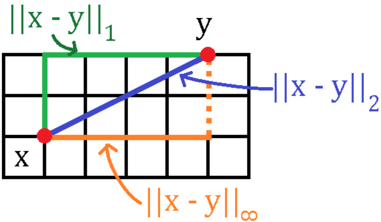
1. $\|\beta x\| = |\beta| \cdot \|x\|$ — Homogeneity
2. $\|x + y\| \leq \|x\| + \|y\|$ — Δ - inequality
3. $\|x\| \geq 0, \ \forall x$ — Non-negativity
4. $\|x\| = 0 \iff x = 0$

Other useful norms

1-norm $\|x\|_1 = |x[1]| + \cdots + |x[n]|$ ("Manhattan metric")

∞-norm $\|x\|_\infty = \max_j |x[j]|$ ("max" norm)

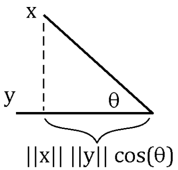
Distance interpretation



Cauchy-Schwartz inequality

Cosine inequality

$$x^T y = \|x\|_2 \|y\|_2 \underbrace{\cos(\theta)}_{\leq 1}$$



Cauchy-Schwartz inequality

$$x^T y \leq \|x\|_2 \|y\|_2$$

with equality of aligned

$$x^T y = \|x\|_2 \|y\|_2 \iff \theta = 0 \iff x = \alpha y$$

for some scalar $\alpha \geq 0$

x.T*y = ||x||_1 ||y||_\infty

Matrix inverses

- $A \in R^{m\times n}$ is the left inverse of $B \in \mathbb{R}^{n\times m}$ if $AB = I$, and the right inverse if $BA = I$
- $A \in \mathbb{R}^{d\times d}$ is the inverse of $B \in \mathbb{R}^{d\times d}$ if $AB = BA = I$ $(A = B^{-1})$
- Left, right, any inverse may not exist

Moore-Penrose pseudoinverse: $A^\dagger$ is a pseudoinverse of $\cancel{B}$ A if

$$AA^\dagger A = A, \quad A^\dagger A A^\dagger = \cancel{I} A^\dagger, \quad (AA^\dagger)^T = AA^\dagger, \quad (A^\dagger A)^T = A^\dagger A$$

1
A = 1 = A dagger
0

Singular value decomposition

The (econo-mode) singular value decomposition (SVD) of a matrix $A \in \mathbb{R}^{m \times n}$ is

$$A = U \Sigma V^T, \quad U \in \mathbb{R}^{m \times r}, \quad \Sigma = \mathbb{R}^{r \times r}, \quad V \in \mathbb{R}^{n \times r}$$

where

m x r r x r r x n

= 0

m x m m x m m x m (if m = n)

- $\underbrace{U^T U = V^T V = I}_{\text{orthonormal}}$ and $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_r)$, $\sigma_i > 0$ for all i
- The columns of U (V) contain left (right) singular vectors of A.
- The diagonal values $\sigma_1, \ldots, \sigma_r$ are the singular values.
- $r \leq \min\{m, n\}$ is the rank of A.

A.T*A*A.T*A

z = randn(d)
z: = A*z/||A*z||_2 --> a million times
z--> top singular vector

s1 = z.T*A*z

Features of SVD

A.T*A
A*A.T --> eigen decomp

{x : x^TA^{-1} x = c}

- If $m = n = r$ then A is invertible; specifically,

$$A^{-1} = U\Sigma^{-1}V^T \text{ and } \Sigma^{-1} = \mathbf{diag}(\sigma_1^{-1}, ..., \sigma_r^{-1})$$

- Otherwise,

$$A^{\dagger} = U\Sigma^{-1}V^T$$

1 1

1 1 I

0

is a valid Moore-Penrose pseudoinverse of A

- If $U = V$ then A is positive semidefinite, e.g.

$$z^T Az \geq 0 \quad \forall z$$

and the SVD is also the eigenvalue decomposition of A

Linear algebra example

Projection on hyperplane

- Given some $\theta \in \mathbb{R}^d$, the set of points in $\mathbb{R}^d$

$$\mathcal{H} = \{x : x^T\theta = 0\}$$

  form a hyperplane

- Given $z \in \mathbb{R}^d$, the Euclidean projection of z onto $\mathcal{H}$ is given by

$$\mathbf{proj}_{\mathcal{H}}(z) = \operatorname*{argmin}_x \{\|x - z\|_2 : x^T\theta = 0\}$$

Linear subspaces

For some matrix $A \in \mathbb{R}^{m \times n}$,

- The range of A is the subspace

$$\mathbf{range}(A) = \{y : Ax = y \text{ for some } x\}.$$

- The nullspace of A is the subspace

$$\mathbf{null}(A) = \{x : Ax = 0\}.$$

- **Decomposition thm**. For any $z \in \mathbb{R}^n$, there exists a unique u, v where

$$z = u + v, \quad u \in \mathbf{null}(A), \quad v \in \mathbf{range}(A^T).$$

Linear decomposition thm. on Euclidean projection

$$\underset{x}{\text{minimize}} \quad \|x - z\|_2 \qquad \text{subject to} \quad x^T\theta = 0$$

Main observation

$$z = \underbrace{z - \frac{z^T\theta}{\theta^T\theta}\theta}_{u} + \underbrace{\frac{z^T\theta}{\theta^T\theta}\theta}_{v}$$

Claim: $u \in \mathbf{null}(\theta^T)$, $v \in \mathbf{range}(\theta)$
Proof:

A
proj_null(A)I - A pinv(A) x
pinv(theta^T) = 1/ (theta^Ttheta) * theta^T

*When we discuss Lagrangian duality, this notion will be made ironclad precise

Linear decomposition thm. on Euclidean projection

$$\underset{x}{\text{minimize}} \quad \|x - z\|_2 \qquad \text{subject to} \quad x^T\theta = 0$$

\in R2

Main observation

$$z = \underbrace{z - \frac{z^T\theta}{\theta^T\theta}\theta}_{u} + \underbrace{\frac{z^T\theta}{\theta^T\theta}\theta}_{v}$$

Claim: $u \in \mathbf{null}(\theta^T)$, $v \in \mathbf{range}(\theta)$
Proof: To show $u \in \mathbf{null}(\theta^T)$, compute

$$\theta^T u = \theta^T z - \frac{z^T\theta}{\theta^T\theta}\theta^T\theta = 0$$

By construction, $v \in \mathbf{range}(\theta)$

*When we discuss Lagrangian duality, this notion will be made ironclad precise

Linear decomposition thm. on Euclidean projection

$$\underset{x}{\text{minimize}} \quad \|x - z\|_2 \qquad \text{subject to} \quad x^T\theta = 0$$

Main observation

$$z = \underbrace{z - \frac{z^T\theta}{\theta^T\theta}\theta}_{u} + \underbrace{\frac{z^T\theta}{\theta^T\theta}\theta}_{v}$$

Claim: $u^T v = 0$
Proof:

*When we discuss Lagrangian duality, this notion will be made ironclad precise

Linear decomposition thm. on Euclidean projection

$$\underset{x}{\text{minimize}} \quad \|x - z\|_2 \qquad \text{subject to} \quad x^T\theta = 0$$

Main observation

$$z = \underbrace{z - \frac{z^T\theta}{\theta^T\theta}\theta}_{u} + \underbrace{\frac{z^T\theta}{\theta^T\theta}\theta}_{v}$$

Claim: $u^T v = 0$
Proof:

$$\begin{aligned} u^T v &= z^T\left(\frac{z^T\theta}{\theta^T\theta}\theta\right) - \left(\frac{z^T\theta}{\theta^T\theta}\right)^2 \theta^T\theta \\ &= \left(\frac{z^T\theta}{\theta^T\theta}\right) z^T\theta - \left(\frac{z^T\theta}{\theta^T\theta}\right)^2 \theta^T\theta \\ &= 0 \end{aligned}$$

*When we discuss Lagrangian duality, this notion will be made ironclad precise

Linear decomposition thm. on Euclidean projection

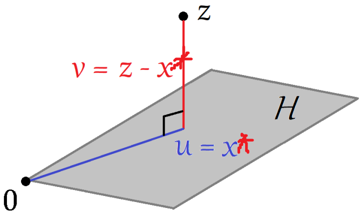
$$\underset{x}{\text{minimize}} \quad \|x - z\|_2 \qquad \text{subject to} \quad x^T\theta = 0$$

Main observation

$$z = \underbrace{z - \frac{z^T\theta}{\theta^T\theta}\theta}_{u} + \underbrace{\frac{z^T\theta}{\theta^T\theta}\theta}_{v}$$

Claim: The minimizer $x^* = u$

Motivational proof by picture:



*When we discuss Lagrangian duality, this notion will be made ironclad precise

Probability and statistics

Discrete distribution

X takes values in a discrete set

- Random variable X has probability mass function (p.m.f.) p_X if

$$p_X(\theta) = \Pr(X = \theta).$$

- A p.m.f. over a finite set Θ is a vector on the unit simplex:

$$p_X(\theta) \geq 0 \; \forall \theta \in \Theta, \quad \sum_{\theta \in \Theta} p_X(\theta) = 1$$

Example: Coin flip. $X =$heads w.p. $1/2$, $X =$ tails w.p. $1/2$

Example: $X = x$ any positive real integer. $p_X(x) = \dfrac{6}{\pi^2} \dfrac{1}{x^2}$.

Continuous distribution

- Random variable X has cumulative distribution function $F_X : \mathbb{R} \to \mathbb{R}$ if

$$F_X(\theta) = \Pr(X \leq \theta).$$

Here,

$$\lim_{\theta \to -\infty} F_X(\theta) = 0, \text{ and } \lim_{\theta \to +\infty} F_X(\theta) = 1$$

and $F_X(\theta)$ is nonzero and monotonically increasing for all θ.

- Derivative of c.d.f. is the probability distribution function (p.d.f.) f_X

$$f_X(\theta) := \frac{d}{d\theta} F_X(\theta)$$

- A valid p.d.f. over Θ must satisfy

$$f_X(\theta) \geq 0 \; \forall \theta \in \Theta, \quad \int_{-\infty}^{\infty} f_X(u) du = 1$$

Distributions

- Gaussian distribution: given mean μ, variance σ,

$$f_X(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- Uniform distribution: Given $a < b$,

$$f_X(\theta) = \begin{cases} \frac{1}{b-a} & a \le \theta \le b \\ 0 & \text{else.} \end{cases}$$

- Spike-and-slab distribution over $\theta \in [0, 1]$: uniform + discrete

$$f_X(\theta) = \begin{cases} w_i \delta_i, & \theta = \theta_i,\ i \in \mathcal{I} \\ \frac{1}{1-\sum_{i \in \mathcal{I}} w_i} & \text{else} \end{cases}$$

- ... many more...

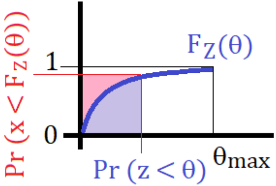
How to sample from a distribution

- Suppose $X \sim$ Unif$(0, 1)$

  (read: X has a uniform p.d.f. with parameters $0 < 1$, or X is distributed uniformly over $0, 1$)

- Then $Z = F_X^{-1}(X)$ has distribution defined by c.d.f. F_X.

- Proof by picture:



$$\Pr(Z < \theta) \overset{\text{def. of c.d.f.}}{=} F_Z(\theta)$$

$$\overset{\text{unif prob.}}{=} \Pr(x < F_Z(\theta))$$

Random vectors

- We say that $X_1, ..., X_d$ are independently identically distributed (i.i.d.) if the distributions are all identical and do not depend on each other:

$$f_X(x_i) = f_X(x_j) = f_X(x_j | x_i) \quad \forall i \neq j$$

- For coupled variables, use random vectors $X \in \mathbb{R}^d$, with nontrivial joint distributions $f_X(x_1, ..., x_d)$
- Example: Gaussian vector parametrized by $\mu \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d \times d}$

$$f_X(x_1, ..., x_d) = \frac{1}{(2\pi)^{d/2}\sqrt{|\Sigma|}} \exp\left((x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

- Question: Are x_i's independent?

Random vectors

- We say that $X_1, ..., X_d$ are independently identically distributed (i.i.d.) if the distributions are all identical and do not depend on each other:

$$f_X(x_i) = f_X(x_j) = f_X(x_j|x_i) \quad \forall i \neq j$$

- For coupled variables, use random vectors $X \in \mathbb{R}^d$, with nontrivial joint distributions $f_X(x_1, ..., x_d)$

- Example: Gaussian vector parametrized by $\mu \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d \times d}$

$$f_X(x_1, ..., x_d) = \frac{1}{(2\pi)^{d/2}\sqrt{|\Sigma|}} \exp\left((x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

- Question: Are x_i's independent?

  Ans: No, unless Σ is diagonal.

Expectations, variance

- The expectation (or mean) of a random variable / vector:

$$\mathbb{E}[X] = \int_{\mathcal{X}} x f_X(x) dx$$

- The variance

$$\mathbf{var}(X) = \mathbb{E}[(X^{\cancel{2}} - \mathbb{E}[X])^2]$$

- Functions of random variables are also random variables
  - Example: $Z = g(X)$
  - Distribution: $p_Z(z) = p_x(g^{-1}(z))$
  - Mean / Expectation: $\mathbb{E}[Z] = \mathbb{E}[g(X)] = \int_{\mathcal{X}} g(x) f_X(x) dx$
  - Variance $\mathbf{var}[Z] = \mathbf{var}[g(X)] = \mathbb{E}[(g(X))^2 - \mathbb{E}[g(X)]^2]$

sample

Empirical statistics

- I want to observe $X \sim f_X$, but I can't see the whole function
- I draw m realizations $x_1, \ldots, x_m \in \mathbb{R}^d$, i.i.d.
- I can guess some things:

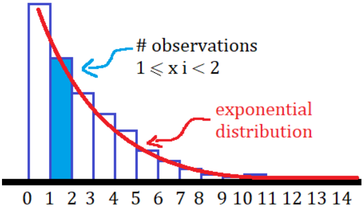
sample mean

$$\text{empirical mean} = \frac{1}{m}\sum_{i=1}^{m} x_i, \qquad \text{empirical covariance} = \frac{1}{m}\sum_{i=1}^{m} x_i x_i^T$$

sample covariance

if mean is 0

- Estimate of p.d.f.: histogram



A couple more useful properties

- Given events $A = a$ and $B = b$,

$$\underbrace{\Pr(A = a, B = b)}_{\text{joint prob.}} = \underbrace{\Pr(A = a | B = b)}_{\text{conditional prob}} \underbrace{\Pr(B = b)}_{\text{marginal prob.}}$$

- Law of Total Probability

$$\Pr(A = a) = \sum_{b \in \mathcal{B}} \Pr(A = a, B = b), \quad \mathcal{B} = \text{ all possible choices of } b$$

- Bayes' rule: glorified rearrangement of first 2 properties

$$\Pr(B = b | A = a) = \frac{\Pr(A = a | B = b) \Pr(B = b)}{\sum_{b' \in \mathcal{B}} \Pr(A = a | B = b') \Pr(B = b')}$$

Bayes' rule example

Bayes' rule is a convenient way of computing likelihoods

$$\Pr(\text{event } k|\text{observation}) = \frac{\Pr(\text{observation}|\text{event } k)\Pr(\text{event } k)}{\sum_i \Pr(\text{observation}|\text{event } i)\Pr(\text{event } i)}$$

Ex. I have 8 red socks and 2 blue socks. I put half my blue socks and all my red socks in the top drawer. The rest I put in the bottom drawer.

I reach and pull a sock from the top drawer. What are the chances that it's blue?

Soln:

$$\Pr(\text{🧦}|\text{top}) = \frac{\Pr(\text{top}|\text{🧦})\Pr(\text{🧦})}{\Pr(\text{top}|\text{🧦})\Pr(\text{🧦}) + \Pr(\text{top}|\text{🧦})\Pr(\text{🧦})}$$

Bayes' rule example

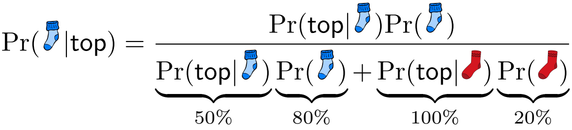
Bayes' rule is a convenient way of computing likelihoods

$$\Pr(\text{event } k|\text{observation}) = \frac{\Pr(\text{observation}|\text{event } k)\Pr(\text{event } k)}{\sum_i \Pr(\text{observation}|\text{event } i)\Pr(\text{event } i)}$$

Ex. I have 8 red socks and 2 blue socks. I put half my blue socks and all my red socks in the top drawer. The rest I put in the bottom drawer.

I reach and pull a sock from the top drawer. What are the chances that it's blue?

Soln:



Bayes' rule example

Bayes' rule is a convenient way of computing likelihoods

$$\Pr(\text{event } k|\text{observation}) = \frac{\Pr(\text{observation}|\text{event } k)\Pr(\text{event } k)}{\sum_i \Pr(\text{observation}|\text{event } i)\Pr(\text{event } i)}$$

Ex. I have 8 red socks and 2 blue socks. I put half my blue socks and all my red socks in the top drawer. The rest I put in the bottom drawer.

I reach and pull a sock from the top drawer. What are the chances that it's blue?

Soln:

$$\Pr({\color{blue}\text{🧦}}|\text{top}) = \frac{\Pr(\text{top}|{\color{blue}\text{🧦}})\Pr({\color{blue}\text{🧦}})}{\underbrace{\Pr(\text{top}|{\color{blue}\text{🧦}})}_{50\%}\underbrace{\Pr({\color{blue}\text{🧦}})}_{80\%} + \underbrace{\Pr(\text{top}|{\color{red}\text{🧦}})}_{100\%}\underbrace{\Pr({\color{red}\text{🧦}})}_{20\%}} = 2/3$$