Fundamentals of Statistics

Unit 0 Review

Normalization constant for the Poisson distribution

1/1 point (graded)

The probability mass function (pmf) of a **Poisson distribution** with parameter λ is given by

$$\mathrm{Poi}\,(\lambda) = \frac{c\lambda^k}{k!}, \qquad k = 0, 1, 2, \dots .$$

Compute the value of c.

$c =$ e^(-lambda) ✔ **Answer:** exp(-lambda)

$e^{-\lambda}$

Solution:

In order to obtain a probability distribution, we must have

$$c \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1. \qquad (1.1)$$

But

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \exp(\lambda) \qquad (1.2)$$

by the series definition of the exponential function. Hence,

$$c = \exp(-\lambda) .$$

Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections: Probability axioms in lecture 1, *Probability models and axioms*.

Moments of Bernoulli variables

3/3 points (graded)

The n^{th} **moment** of a random variable X is defined to be the expectation $\mathbb{E}[X^n]$ of the n^{th} power of X.

Recall that a **Bernoulli random variable with parameter p** is a random variable that takes the value 1 with probability p, and the value 0 with probability $1-p$.

Let X be a Bernoulli random variable with parameter 0.7. Compute the **expectation values** of X^k, denoted by $\mathbb{E}[X^k]$, for the following three values of k: $k = 1, 4$, and 3203.

$\mathbb{E}[X] =$ 0.7 ✔ **Answer:** 0.7

$\mathbb{E}[X^4] =$ 0.7 ✔ **Answer:** 0.7

$\mathbb{E}[X^{3203}] =$ 0.7 ✔ **Answer:** 0.7

Solution:

Remember, the expectation of a discrete random variable is

$$\mathbb{E}[X] = \sum_{j \in \text{range}(X)} j\,\mathbf{P}(X = j),$$

while the higher moments are

$$\mathbb{E}[X^n] = \sum_{j \in \text{range}(X)} j^n\,\mathbf{P}(X = j),$$

For a Bernoulli random variable with parameter p, the range is $\{0, 1\}$, and $0^k = 0$, $1^k = 1$ for all $k \geq 1$, so all moments are equal to the first one,

$$\mathbb{E}[X] = 0 \times (1-p) + 1 \times p = p,$$

and we get the result by plugging in $p = 0.7$.
Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections: Expectation in lecture 5, *Probability mass functions and expectations*.

Variance of Bernoulli variables

3/3 points (graded)

Let X be a Bernoulli random variable with parameter $p \in [0,1]$. Compute the **variance** of X, which is denoted by $\mathsf{Var}[X]$.

$\mathsf{Var}[X] =$ p*(1-p) ✔ **Answer:** p*(1-p)

$p \cdot (1-p)$

What value(s) of the parameter p maximize the variance? What values minimize it?

(For each question, enter the values of p as a list of **numbers**, separated by commas. For example, to enter the set $\{0.2, 0.3\}$, type **0.2, 0.3**, **without the braces**. Do NOT enter duplicate values, e.g. **0.2, 0.3, 0.3** will be graded as incorrect.)

The values of p for which $\mathsf{Var}[X]$ is minimized: 0,1 ✔ **Answer:** 0, 1

$0, 1$

The values of p for which $\mathsf{Var}[X]$ is maximized: 0.5 ✔ **Answer:** 1/2

0.5

Solution:

Recall from the previous exercise that $\mathbb{E}[X^n] = p$ for all positive integers n. Therefore, the variance is

$$\mathsf{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1-p).$$

This is a quadratic polynomial with negative leading factor, hence it does not attain a global minimum on $\mathbb{R}$. For the range $p \in [0,1]$ in question, its minima are attained at both boundary points $p = 0$ and $p = 1$. Its maximum can be found by differentiating and setting the derivative equal to zero. It occurs at $p = \frac{1}{2}$.

Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections: Variance in lecture 6, *Variance; conditioning on an event; multiple random variable*.

Sum of Bernoulli variables

1/1 point (graded)

Given n i.i.d. realizations $X_1, \ldots, X_n \sim \mathrm{Ber}(p)$, what is the distribution of $\sum_{i=1}^n X_i$?

- Poisson with parameter pn
- Gamma with parameters n and p
- Binomial with parameters n and p
- Bernoulli with parameter pn

STANDARD NOTATION

Solution:

We know from probability theory that $\sum_{i=1}^n X_i$ follows a Binomial distribution with parameters n and p.

Discrete uniform random variables

2/2 points (graded)

Recall that a **uniform random variable** is a random variable that takes values with equal probability,

Let X be a uniform random variable in the finite set $\{1, 2, \ldots, 20\}$.

Compute the following quantities.

The probability that X is an even number:

$\mathbf{P}\left(X \text{ is an even number}\right) =$ 1/2 ✔ **Answer:** 1/2

The probability that X is a prime number:

$\mathbf{P}\left(X \text{ is a prime number}\right) =$ 8/20 ✔ **Answer:** 2/5

STANDARD NOTATION

Solution:

There are 10 even numbers in $\{1, \ldots, 20\}$, therefore

$$\mathbf{P}\left(X \text{ is an even number}\right) = \frac{10}{20} = \frac{1}{2}.$$

There are 8 prime numbers in $\{1, \ldots, 20\}$, (namely $\{2, 3, 5, 7, 11, 13, 17, 19\}$, so

$$\mathbf{P}\left(X \text{ is a prime number}\right) = \frac{8}{20} = \frac{2}{5}.$$

Moments - http://www.milefoot.com/math/stat/rv-moments.htm

Moments of Gaussian random variables

2/5 points (graded)

Let X be a Gaussian random variable with mean μ and variance σ^2. Compute the following moments:

Remember that we use the terms **Gaussian random variable** and **normal random variable** interchangeably.

(Enter your answers in terms of μ and σ.)

$\mathbb{E}[X^2] =$ sigma^2 + mu^2 ✔ **Answer:** sigma^2 + mu^2

$\sigma^2 + \mu^2$

$\mathbb{E}[X^3] =$ 3*mu*(sigma^2+mu^2 ✔ **Answer:** 3 * sigma^2 * mu + mu^3

$3 \cdot \mu \cdot (\sigma^2 + \mu^2) - 2 \cdot \mu^3$

$\mathbb{E}[X^4] =$ (12*mu^2*(sigma^2+m ✘ **Answer:** 3 * sigma^4 + 6 * sigma^2 * mu^2 + mu^4

$(12 \cdot \mu^2 \cdot (\sigma^2 + \mu^2) - 8 \cdot \mu^4) - (6 \cdot \mu^2 \cdot \sigma^2 + 6 \cdot \mu^4) + 3 \cdot \mu^4$

Var $(X^2) =$ (sigma^2 + mu^2)^2 - ((✘ **Answer:** 2 * sigma^4 + 4 * sigma^2 * mu^2

$(\sigma^2 + \mu^2)^2 - (4 \cdot \mu \cdot (3 \cdot \mu \cdot (\sigma^2 + \mu^2) - 2 \cdot \mu^3) - 6 \cdot \mu^2 \cdot (\sigma^2 + \mu^2) + 3 \cdot \mu^4)$

Write $\mathbf{P}(X > 0)$ in terms of the **cumulative distribution function (cdf)** Φ of the standard Gaussian distribution, evaluated at a function of μ and σ.

Recall that

$$\Phi(x) = \mathbf{P}(Z \leq x), \quad x \in \mathbb{R},$$

where $Z \sim \mathcal{N}(0, 1)$ is a standard normal variable. (Enter `Phi` for Φ, e.g. if the answer is $\Phi(\mu)$, enter `Phi(mu)`.)

$\mathbf{P}(X > 0) =$ Phi(0) ✘ **Answer:** 1 - Phi(-mu/sigma)

$\Phi(0)$

Solution:

We can write a general Gaussian variable $X \sim \mathcal{N}(\mu, \sigma^2)$ as $X = \sigma Z + \mu$, where $Z \sim \mathcal{N}(0, 1)$ is a standard normal variable. Hence, the calculation can be made by factoring out the corresponding polynomials and calculating (or looking up) the moments of Z:

$$\begin{aligned}\mathbb{E}[Z] &= 0\\ \mathbb{E}[Z^2] &= 1\\ \mathbb{E}[Z^3] &= 0\\ \mathbb{E}[Z^4] &= 3.\end{aligned}$$

As an example, let us compute $\mathbb{E}[X^3]$. Denote the density of a standard normal distribution by $\varphi(z)$, i.e.,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

With this, we calculate

$$\begin{aligned}\mathbb{E}[X^3] &= \int_{-\infty}^{\infty} (\sigma z + \mu)^3 \phi(z)\, dz\\ &= \sigma^3 \mathbb{E}[Z^3] + 3\sigma^2 \mu \mathbb{E}[Z^2] + 3\sigma\mu^2 \mathbb{E}[Z] + \mu^3\\ &= 3\sigma^2\mu + \mu^3.\end{aligned}$$

For $\text{Var}(X^2)$, we can use the formula $\text{Var}(X^2) = \mathbb{E}[X^4] - \left(\mathbb{E}[X^2]\right)^2$.

Similarly, we can express the probability $\mathbf{P}(X > 0)$ as

$$\begin{aligned}\mathbf{P}(X > 0) &= \mathbf{P}(\sigma Z + \mu > 0) = \mathbf{P}(\sigma Z > -\mu)\\ &= \mathbf{P}(Z > -\frac{\mu}{\sigma}) = 1 - \Phi\left(-\frac{\mu}{\sigma}\right).\end{aligned}$$

Covariance of Gaussians

4/4 points (graded)

Recall that **i.i.d.** stands for **independent and identically distributed** . A collection of random variables $X_1, \ldots, X_n$ are **i.i.d.** if all of them follow the same distribution, and each X_i does not contain information about the other realizations.

Let X, Y be i.i.d. **standard** normal random variables, that is, $X, Y \sim \mathcal{N}(0, 1)$.

Recall that the **covariance** of two random variables X and Y, denoted by $\text{Cov}(X, Y)$, is defined as

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]. \tag{1.3}$$

Compute the following variances and covariances.

$\text{Var}(X + Y) =$ 2

$\text{Var}(XY) =$ 1

$\text{Cov}(X, X + Y) =$ 1

$\text{Cov}(X, XY) =$ 0

Solution:

Note that by the definition of a standard Gaussian random variable,

$$\mathbb{E}[X] = \mathbb{E}[Y] = 0 \quad \mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1.$$

With this, compute

$$\begin{aligned} \mathsf{Var}(X+Y) = &\ \mathsf{Var}(X) + \mathsf{Var}(Y) && \text{(independence)} \\ = &\ 1+1=2, \end{aligned}$$

$$\begin{aligned} \mathsf{Var}(XY) = &\ \mathbb{E}[(XY)^2] - (\mathbb{E}[XY])^2 \\ = &\ \mathbb{E}[X^2]\,\mathbb{E}[Y^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2 && \text{(independence)} \\ = &\ 1 \times 1 - 0 = 1, \end{aligned}$$

$$\begin{aligned} \mathsf{Cov}(X, X+Y) = &\ \mathbb{E}[X(X+Y)] - \mathbb{E}[X]\,\mathbb{E}[X+Y] \\ = &\ \mathbb{E}[X^2] + \mathbb{E}[XY] - \mathbb{E}[X](\mathbb{E}[X] + \mathbb{E}[Y]) && \text{(linearity of expectation)} \\ = &\ \mathbb{E}[X^2] + \mathbb{E}[X]\,\mathbb{E}[Y] - \mathbb{E}[X]^2 - \mathbb{E}[X]\,\mathbb{E}[Y] && \text{(independence)} \\ = &\ 1, \end{aligned}$$

$$\begin{aligned} \mathsf{Cov}(X, XY) = &\ \mathbb{E}[X(XY)] - \mathbb{E}[X]\,\mathbb{E}[XY] \\ = &\ \mathbb{E}[X^2]\,\mathbb{E}[Y] - \mathbb{E}[X]^2\mathbb{E}[Y] && \text{(independence)} \\ = &\ 1 \cdot 0 - 0 \cdot 0 = 0. \end{aligned}$$

Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections: 8. Covariance, 9. Covariance properties, and 10. the variance of a sum in Lecture 12, *Sums of independent random variables; covariance, and correlation*.

True or False: Variance, covariance and independence

2/2 points (graded)

For each of the statements below, determine whether it is true (meaning, always true) or false (meaning, not always true).

- For any two random variables, $\mathsf{Var}(X+Y) = \mathsf{Var}(X) + \mathsf{Var}(Y)$.

  - True
  - **False** ✔

- If the covariance, $\mathrm{Cov}(X, Y)$ between two random variables X, Y is 0, then X and Y are independent.

  - True
  - **False** ✔

Solution:

- The first item is False. For any two random variables, it is known that,

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y).$$

In particular, if $\text{Cov}(X,Y) \neq 0$, this does not hold.

- The second item is also false. As a simple example, let $X \sim \text{Unif}[-1,1]$ and let $Y = X^2$. Then,

$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2] = 0,$$

using the fact that X is centered and symmetric around 0, and its odd moments vanish. Even though they are uncorrelated, they are (highly) dependent, Y is obtained from X, intuitively!

4. Uniform random variables

Bookmark this page

Expectation, variance and probabilities

4/4 points (graded)

Let X be a uniform random variable in the interval $[2, 8.5]$. Find the following quantities (if needed, round to the nearest 10^{-4}):

$\mathbb{E}[X] =$ 5.25 ✔ **Answer:** 21/4

$\text{Var}[X] =$ 3.5208 ✔ **Answer:** 169/48

$\mathbf{P}(X > 4) =$ 0.6923 ✔ **Answer:** 9/13

$\mathbf{P}(\log(X) \leq 1) =$ 0.1105 ✔ **Answer:** 0.11054897

Solution:

We can write $X = 6.5Z + 2$, where Z follows a uniform distribution on $[0, 1]$. By properties of the uniform distribution, we then conclude:

$$\mathbb{E}[X] = 2 + 6.5\mathbb{E}[Z] = 2 + 6.5 \times \frac{1}{2} = \frac{21}{4},$$

$$\text{Var}[X] = 6.5^2 \times \text{Var}[Z] = \frac{169}{4} \times \frac{1}{12} = \frac{169}{48},$$

$$\mathbf{P}(X \geq 4) = \mathbf{P}(Z \geq \frac{4}{13}) = 1 - \frac{4}{13} = \frac{9}{13},$$

$$\mathbf{P}(\log(X) \leq 1) = \mathbf{P}(X \leq e) = \mathbf{P}(Z \leq \frac{2(\mathbf{e} - 2)}{13}) \approx 0.110549.$$

Two independent copies

2/3 points (graded)

Let U, V be i.i.d. random variables uniformly distributed in $[0,1]$. Compute the following quantities:

$\mathbb{E}\left[\,|U-V|\,\right] =$ 0.25 ✖ **Answer:** 1/3

$\mathbf{P}(U=V) =$ 0 ✔ **Answer:** 0

$\mathbf{P}(U \le V) =$ 0.5 ✔ **Answer:** 1/2

STANDARD NOTATION

Solution:

For the first quantity, we write the joint expectation as an iterated expectation and conditional expectation,

$$\mathbb{E}\left[|U-V|\right] = \mathbb{E}\left[\mathbb{E}\left[|U-V|\,|V\right]\right].$$

By independence, we can compute the inner expectation as

$$\begin{aligned}
\mathbb{E}\left[|U-V|\,|V=v\right] &= \int_0^1 |u-v|\,du \\
&= \int_0^v (v-u)\;du + \int_v^1 (u-v)\;du \\
&= \left[vu - \frac{1}{2}u^2\right]_0^v + \left[\frac{1}{2}u^2 - vu\right]_v^1 = v^2 - \frac{1}{2}v^2 + \frac{1}{2} - v - \frac{1}{2}v^2 + v^2 \\
&= v^2 - v + \frac{1}{2},
\end{aligned}$$

Generating Speech Output

so

$$\mathbb{E}\left[\,|U-V|\,\right] = \mathbb{E}\left[V^2 - V + \frac{1}{2}\right] = \frac{1}{3} - \frac{1}{2} + \frac{1}{2} = \frac{1}{3}.$$

For the probability $\mathbf{P}(U = V)$, just write this as double expectation as well and notice that

$$\mathbf{P}(U = V) = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}(U = V)\,|V\right]\right] = \mathbb{E}[0] = 0,$$

because the probability of a uniform random variable being equal to any fixed number between 0 and 1 is zero.

For $\mathbf{P}(U \leq V)$, write it again as a double expectation,

$$\mathbf{P}(U \leq V) = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}(U \leq V)\,|V\right]\right] = \mathbb{E}\left[\mathbf{P}(U \leq V)\,|V\right] = \mathbb{E}[V] = \frac{1}{2}.$$

Alternatively, this can also be seen by symmetry of the two variables, i.e., $P(U \leq V) = P(V \leq U)$ and either one of the two must be true, counting double the zero-set of $\mathbf{P}(U = V)$.
Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections: Uniform PDF in Lecture 8, *Probability density functions*.

Maximum and sum of independent copies

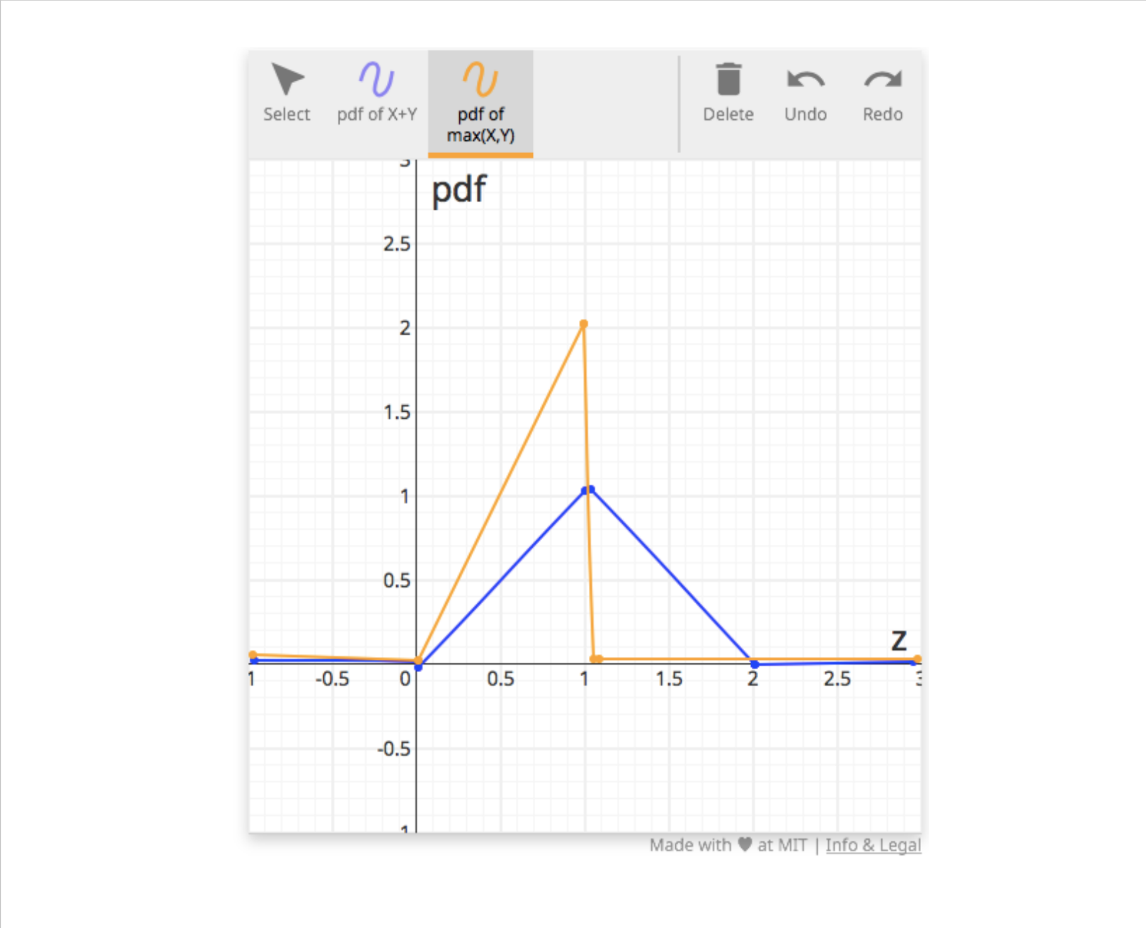
0/1 point (graded)

Let X, Y be independent random variables uniformly distributed in $[0,1]$. In the graph below, sketch

1. the probability density $f_{X+Y}(z)$ of $X+Y$;
2. the probability density $f_{\max(X,Y)}(z)$ of $\max(X,Y)$.

(Be sure to sketch on the **entire domain** shown on the graph.)

Drawing tip: The spline tool draws a smooth curve connecting the points you click. To draw sharp corners, click on the point where the corner would be, then click again very close to it, and then continue onto the next point of your function.



The density of $X + Y$ is given by the convolution of the density of a uniform random variable,

$$f(x) = \mathbf{1}[0,1] = \begin{cases} 1, & \text{if } x \in [0,1] \\ 0, & \text{otherwise} \end{cases}$$

The density of $X+Y$ is given by the convolution of the density of a uniform random variable,

$$f(x) = \mathbf{1}[0,1] = \begin{cases} 1, & \text{if } x \in [0,1] \\ 0, & \text{otherwise} \end{cases}$$

The density g of $X+Y$ therefore is

$$\begin{aligned} g(z) = & \int_{\mathbb{R}} f(x) f(z-x)\, dx \\ = & \int_{\mathbb{R}} \mathbf{1}(x \in [0,1])\, \mathbf{1}(z - x \in [0,1])\, dx \\ = & \int_0^1 \mathbf{1}(z-1 \le x \le z)\, dx \\ = & \ \mathbf{1}(z \le 2) \int_{\max\{0,z-1\}}^{\min\{1,z\}} dx \\ = & \begin{cases} 0, & z < 0 \\ z, & 0 < z < 1 \\ 2-z, & 1 < z < 2 \\ 0, & z > 2. \end{cases} \end{aligned}$$

For the density of $\max\{X,Y\}$, first note that it is supported in $[0,1]$. Now, first compute the cdf on that interval:

$$\begin{aligned} \mathbf{P}(\max\{X,Y\} \le y) = & \ \mathbf{P}(X \le y)\, \mathbf{P}(Y \le y) \quad \text{(by independence)} \\ = & \ t^2. \end{aligned}$$

Hence, the density h of $\max\{X,Y\}$ is given by

$$h(z) = \begin{cases} 0, & z < 0 \\ 2z, & 0 \le z \le 1 \\ 0, & z > 1. \end{cases}$$

Maximum of uniform random variables

1/2 points (graded)

Let $U_1, \ldots, U_n$ be i.i.d. random variables uniformly distributed in $[0, 1]$ and let $M_n = \max_{1 \le i \le n} U_i$.

Find the cdf of M_n, which we denote by $G(t)$, for $t \in [0, 1]$.

For $t \in [0, 1]$,

$G(t) =$ t^n ✔ **Answer:** t^n

t^n

Now, let $F_n(t)$ denote the cdf of $n(1 - M_n)$; for $t > 0$, compute

$\lim_{n \to \infty} F_n(t) =$ 0 ✖ **Answer:** 1-exp(-t)

0

Solution:

First, we compute the cdf. Let $t \in [0, 1]$. Then,

$$\mathbf{P}(M_n \le t) = \mathbf{P}\left(\max_{i=1,\ldots,n} U_i \le t\right) = \mathbf{P}\left(\cap_{i=1}^n \{U_i \le t\}\right) = \prod_{i=1}^n \mathbf{P}(U_i \le t) = t^n,$$

where we used the independence of the U_i to write the intersection as a product.

Now,

$$\begin{aligned}\mathbf{P}(n(1 - M_n) \le t) = &\ \mathbf{P}\left(1 - M_n \le \frac{t}{n}\right) = \mathbf{P}\left(M_n \ge 1 - \frac{t}{n}\right) \\ = &\ 1 - \mathbf{P}\left(M_n < 1 - \frac{t}{n}\right) = 1 - \left(1 - \frac{t}{n}\right)^n \xrightarrow{n \to \infty} 1 - \mathbf{e}^{-t}.\end{aligned}$$

Hence, $n(1 - M_n)$ converges in distribution to $\text{Exp}(1)$.

5. Exponential random variables

Bookmark this page

Sums and products

1/3 points (graded)

Let X be an exponential random variable with parameter $\lambda > 0$ and Y be a Poisson random variable with parameter $\mu > 0$. Assume that X and Y are independent. Compute the following quantities:

$\mathbb{E}[X^2 + Y^2] =$ (2/lambda^2) + 1/mu^2 ✖ **Answer:** 2/(lambda^2)+ mu + mu^2

$\left(\frac{2}{\lambda^2}\right) + \frac{1}{\mu^2}$

$\mathbb{E}[X^2 Y] =$ (2/lambda^2) *mu ✔ **Answer:** 2 * mu / (lambda^2)

$\left(\frac{2}{\lambda^2}\right) \cdot \mu$

$\text{Var}(2X + 3Y) =$ (2/lambda) + 3*mu ✖ **Answer:** 4/(lambda^2) + 9 * mu

$\left(\frac{2}{\lambda}\right) + 3 \cdot \mu$

Solution:

First, let us review the moments of the Exponential and Poisson distribution:

If $X \sim \text{Exp}(\lambda)$ with $\lambda > 0$, then

$$\mathbb{E}[X] = \lambda, \quad \mathbb{E}[X^2] = \frac{2}{\lambda^2}, \quad \text{Var}(X) = \frac{1}{\lambda^2}. \tag{1.4}$$

If $Y \sim \text{Poi}(\mu)$, again with $\mu > 0$, then

$$\mathbb{E}[Y] = \mu, \quad \mathbb{E}[Y^2] = \mu + \mu^2, \quad \text{Var}(Y) = \mu. \tag{1.5}$$

Now, we can use the rules for expectation and variance to calculate:

$$\begin{aligned}
\mathbb{E}[X^2 + Y^2] &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] && \text{(linearity of expectation)}\\
&= \frac{2}{\lambda^2} + \mu + \mu^2 \\
\mathbb{E}[X^2 Y] &= \mathbb{E}[X^2]\,\mathbb{E}[Y] && \text{(multiplicativity of expectation for independent vari}\\
&= \frac{2\mu}{\lambda^2} \\
\text{Var}(2X + 3Y) &= \text{Var}(2X) + \text{Var}(3Y) && \text{(additivity of variance for independent variables)}\\
&= 2^2\text{Var}(X) + 3^2\text{Var}(Y) && \text{(scaling property of variance)}\\
&= \frac{4}{\lambda^2} + 9\mu
\end{aligned}$$

Estimators

0/1 point (graded)

Let $X_1, \ldots, X_n$ be i.i.d exponential random variables with parameter λ and let $Z_i = \mathbf{1}(X_i \leq 1), i = 1, \ldots, n$. Recall that $\mathbf{1}(X \leq 1)$ denotes the **indicator function** that takes the value 1 when $X \leq 1$ and 0 otherwise.

What is the limit in probability, as n goes to infinity, of $\frac{1}{n}\sum_{i=1}^{n} Z_i$?

$\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow[n\to\infty]{\mathbf{P}}$ 1

1

✖ **Answer:** 1 - exp(-lambda)

Solution:

Since the X_i are independent and identically distributed, so are the Z_i. By the Law of Large Numbers, we know that

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[Z_i],$$

so it is enough to compute that quantity.

For this, note that

$$\mathbb{E}[Z_i] = \mathbf{P}(X_i \leq 1) = 1 - \exp(-\lambda \times 1) = 1 - \exp(-\lambda),$$

which follows from the formula for the cdf of an Exponential distribution. Hence,

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow[n\to\infty]{\mathbf{P}} 1 - \exp(-\lambda),$$

Refer to *6.431x Probability–the Science of Uncertainty and Data*. Sections:

Properties of the exponential distribution

2/2 points (graded)

Let X be an exponential random variable with parameter $\lambda = 2$ that models the lifetime (in years) of a lightbulb. Compute the probability that the lightbulb lasts for at least 2 years. Round your answer to the nearest 10^{-2}.

$\mathbf{P}(X \geq 2) =$ 0.0183 ✔ **Answer:** exp(-4)

Given the lightbulb has lasted 2 years, find the probability that it lasts for k more years for any positive integer k.

$\mathbf{P}(X \geq k+2 | X \geq 2) =$ exp(-2*k) ✔ **Answer:** exp(-2*k)

$\exp(-2 \cdot k)$

Solution:

The exponential distribution with parameter λ has a continuous density on $(0, \infty)$ with cdf

$$F(x) = 1 - \exp(-\lambda x).$$

Hence, for $\lambda = 2$,

$$\mathbf{P}(X \geq 2) = 1 - \mathbf{P}(X \leq 2) = 1 - (1 - \exp(-2 \times 2)) = e^{-4}.$$

For the second part, note that $\{X \geq k+2\} \subseteq \{X \geq 2\}$. Therefore,

$$\mathbf{P}(X \geq k+2 | X \geq 2) = \frac{\mathbf{P}(\{X \geq k+2\} \cap \{X \geq 2\})}{\mathbf{P}(X \geq 2)} = \frac{\mathbf{P}(X \geq k+2)}{\mathbf{P}(X \geq 2)} = \frac{e^{-2(k+2)}}{e^{-4}} = e^{-2k}.$$

Remark: This is an example of the exponential distribution being **memoryless** : The probability of the lightulb lasting k more years given that it already lasted 2 years is exactly the same as the probability of it lasting k years in the first place.

Gaussian probabilities

1/4 points (graded)

Let $X \sim N(1, 2.25)$. As a reminder, the 2.25 here represents the value of σ^2. Using the normal probability table below, compute the following probabilities:

Normal probability table

Show

$\mathbf{P}(X > 1) =$ 0.5 ✔

$\mathbf{P}(|X - 2| \leq 1) =$ 0.9082 ✖

$\mathbf{P}(X^2 > 4) =$ 0.5092 ✖

$\mathbf{P}(X^2 - 2X - 1 > 0) =$ 0.1736 ✖

Solution:

First, note that for $Z \sim \mathcal{N}(0, 1)$, $x > 0$, we have

$$\mathbf{P}(Z \leq -x) = \mathbf{P}(Z \geq x) = 1 - \mathbf{P}(Z \leq x),$$

and

$$\mathbf{P}(Z \geq x) = 1 - \mathbf{P}(Z \leq x).$$

Moreover, if $X \sim \mathcal{N}(1, 2.25)$, we can write it as $X = 1.5Z + 1$, where $Z \sim \mathcal{N}(0, 1)$. This allows us to reduce all probabilities to the ones that are listed in the table.

In particular,

$$\begin{aligned}
\mathbf{P}(X > 1) = {} & \mathbf{P}(1.5Z + 1 > 1) = \mathbf{P}(1.5Z > 0) \\
= {} & \mathbf{P}(Z \geq 0) = 1 - \mathbf{P}(Z \leq 0) = 1 - 0.5000 = 0.5000,
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(|X - 2| \leq 1) = {} & \mathbf{P}(-1 \leq (X - 2) \leq 1) = \mathbf{P}(-1 \leq (1.5Z + 1 - 2) \leq 1) \\
= {} & \mathbf{P}(0 \leq 1.5Z \leq 2) \\
\simeq {} & \mathbf{P}(0 \leq Z \leq 1.33) \\
= {} & \mathbf{P}(Z \leq 1.33) - \mathbf{P}(Z \leq 0) \simeq 0.9082 - 0.5000 = 0.4082
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(X^2 > 4) = {} & \mathbf{P}(|X| > 2) = \mathbf{P}(|1.5Z + 1| > 2) \\
= {} & \mathbf{P}(1.5Z + 1 \leq -2) + \mathbf{P}(1.5Z + 1 \geq 2) \\
= {} & \mathbf{P}(Z \leq -2) + \mathbf{P}\left(Z \geq \frac{2}{3}\right) \\
= {} & 1 - \mathbf{P}(Z \leq 2) + 1 - \mathbf{P}\left(Z \leq \frac{2}{3}\right) \\
\simeq {} & 2 - 0.9772 - 0.7486 = 0.2742
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(X^2 - 2X - 1 > 0) = {} & \mathbf{P}((X - 1)^2 - 2 > 0) = \mathbf{P}(|X - 1| > \sqrt{2}) \\
= {} & \mathbf{P}(|1.5Z| > \sqrt{2}) \\
= {} & \mathbf{P}(Z > \frac{\sqrt{2}}{1.5}) + \mathbf{P}(Z < -\frac{\sqrt{2}}{1.5}) \\
= {} & 2 - 2\mathbf{P}(Z < \frac{\sqrt{2}}{1.5}) \\
\simeq {} & 2 - 2\,(0.8264) \\
= {} & 0.3472.
\end{aligned}$$

Approximation of Binomial variables

1/1 point (graded)

Using the normal probability table, evaluate approximately $\mathbf{P}(X > 400)$, where X is a binomial random variable with parameters 1000 and $.3$.

Normal probability table

Show

$\mathbf{P}(X > 400) \simeq$ 0.0001 ✔ **Answer:** 0.0002

STANDARD NOTATION

Solution:

A binomial distribution with parameters (n, p) has expectation np and variance $np(1-p)$. Hence, by the Central Limit Theorem, we have

$$\frac{1}{\sqrt{np(1-p)}}(X - np) \xrightarrow{\text{(D)}} Z \sim \mathcal{N}(0, 1).$$

The probability in question can therefore be approximated by

$$\begin{aligned}
\mathbf{P}(X > 400) = \;& \mathbf{P}\left(\frac{1}{\sqrt{1000 \times 0.3 \times 0.7}}(X - 300) > \frac{100}{\sqrt{1000 \times 0.3 \times 0.7}}\right) \\
\simeq \;& 1 - \mathbf{P}(Z \leq \frac{100}{\sqrt{1000 \times 0.3 \times 0.7}}) \\
\simeq \;& 1 - \mathbf{P}(Z \leq 6.90) \\
\leq \;& 1 - 0.9998 = 0.0002.
\end{aligned}$$

Matrix Multiplication

6/6 points (graded)

Let $\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & -4 \end{pmatrix}$ and let $\mathbf{B} = \begin{pmatrix} -1 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}$. The dimensions of the product $\mathbf{AB}$ are:

2 ✔ **Answer:** 2 rows × 3 ✔ **Answer:** 3 columns.

More generally, let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times k$ matrix. What is the size of $\mathbf{AB}$?

m ✔ **Answer:** m rows × k ✔ **Answer:** k columns.

In addition, if $\mathbf{C}$ is a $k \times j$ matrix, what is the size of $\mathbf{ABC}$?

m ✔ **Answer:** m rows × j ✔ **Answer:** j columns.

Solution:

The size of the output is the number of rows of the left matrix, and the number of columns of the right matrix. The two dimensions on the inside (columns of the left matrix, rows of the right matrix) must match.
In the first part, $\mathbf{AB}$ is 2×3.
For the second and third parts, $\mathbf{AB}$ is $m \times k$ and $\mathbf{ABC}$ is $m \times j$.

Vector Inner product

1/1 point (graded)

Suppose $\mathbf{u} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. The product $\mathbf{u}^T\mathbf{v}$ evaluates the **inner product** (also called the **dot product**) of $\mathbf{u}$ and $\mathbf{v}$, which evaluates to

$\mathbf{u}^T\mathbf{v} =$ 2 ✔ **Answer:** 2

The inner product of $\mathbf{u}$ and $\mathbf{v}$ is sometimes written as $\langle \mathbf{u}, \mathbf{v} \rangle$.

Solution:

The inner product is always a scalar (a 1×1 matrix). In this case, it evaluates to $1 \cdot -1 + 3 \cdot 1 = 2$. In general, if $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$, then $\mathbf{u}^T\mathbf{v} = \sum_{i=1}^{n} u_i v_i$.

$$\begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (\cdot)$$

Vector Outer product

4/4 points (graded)

Suppose $\mathbf{u} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. The product $\mathbf{u}\mathbf{v}^T$ evaluates the **outer product** of $\mathbf{u}$ and $\mathbf{v}$, which is a 2×2 matrix in this case.

What is $(\mathbf{u}\mathbf{v}^T)_{1,1}$?

-1

✔ **Answer:** -1

What is $(\mathbf{u}\mathbf{v}^T)_{1,2}$?

1

✔ **Answer:** 1

What is $(\mathbf{u}\mathbf{v}^T)_{2,1}$?

-3

✔ **Answer:** -3

What is $(\mathbf{u}\mathbf{v}^T)_{2,2}$?

3

✔ **Answer:** 3

Solution:

In this case, the outer product evaluates to

$$\mathbf{u}\mathbf{v}^T = \begin{pmatrix} -1 & 1 \\ -3 & 3 \end{pmatrix}.$$

In general, if $\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, $\mathbf{u}\mathbf{v}^T$ is an $m \times n$ matrix whose (i, j) entry is $(\mathbf{u}\mathbf{v}^T)_{i,j} = u_i v_j$.

Matrices - rank https://stattrek.com/matrix-algebra/matrix-rank.aspx

8. Linear Independence, Subspaces and Dimension

Bookmark this page

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are said to be **linearly dependent** if there exist scalars $c_1, \ldots, c_n$ such that (1) not all c_i's are zero and (2) $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = 0$.

Otherwise, they are said to be **linearly independent** : the only scalars $c_1, \ldots, c_n$ that satisfy $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = 0$ are $c_1 = \cdots = c_n = 0$.

The collection of non-zero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m$ determines a **subspace** of $\mathbb{R}^m$, which is the set of all linear combinations $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ over different choices of $c_1, \ldots, c_n \in \mathbb{R}$. The **dimension** of this subspace is the size of the **largest possible, linearly independent** sub-collection of the (non-zero) vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Row and Column Rank

2/2 points (graded)

Suppose $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$. The rows of the matrix, $(1, 3)$ and $(2, 6)$, span a subspace of dimension

1

Answer: 1 . This is the **row rank** of $\mathbf{A}$.

The columns of the matrix, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$ span a subspace of dimension

1

Answer: 1 . This is the **column rank** of $\mathbf{A}$.

We will be using these ideas when studying **Linear Regression**, where we will work with larger, possibly rectangular matrices.

Solution:

In both cases, the two vectors are linearly dependent.

$$2 \cdot (1, 3) - (2, 6) = (0, 0)$$

$$3\begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The rank of a matrix

3/3 points (graded)

In general, row rank is always equal to the column rank, so we simply refer to this common value as the **rank** of a matrix.

What is the largest possible rank of a 2×2 matrix?

2

✔ **Answer:** 2

What is the largest possible rank of a 5×2 matrix?

2

✔ **Answer:** 2

In general, what is the largest possible rank of an $m \times n$ matrix?

- ○ m
- ○ n
- ◉ $\min(m, n)$
- ○ $\max(m, n)$
- ○ None of the above

✔

Solution:

In general, the rank of any $m \times n$ matrix can be at most $\min(m, n)$, since rank = column rank = row rank. For example, if there are five columns and three rows, the column rank cannot be larger than the largest possible row rank – the largest possible row rank for three rows is, unsurprisingly, 3. The opposite is also true if there are more rows than columns. If a matrix has two columns and six rows, then the row rank cannot exceed the column rank, which is at most 2.
In general, a matrix $\mathbf{A}$ is said to have **full rank** if $\operatorname{rank}(\mathbf{A}) = \min(m, n)$. (note the $=$, instead of $\leq$).

Examples of rank

5/5 points (graded)

What is the rank of $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$?

1

✔ **Answer:** 1

What is the rank of $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$?

2

✔ **Answer:** 2

What is the rank of $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$?

0

✔ **Answer:** 0

What is the rank of $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$?

2

✔ **Answer:** 2

What is the rank of $\begin{pmatrix} 1 & 1 & 0 \\ 0 & -3 & 2 \\ 0 & 0 & 1 \end{pmatrix}$?

3

✔ **Answer:** 3

Solution:

1. The set of rows describe a subspace of dimension 1, spanned by $(1, 1)$.

2. This matrix has rank 2, since $(1, -1)$ and $(1, 0)$ are linearly independent.

3. This matrix has rank zero. By definition, the rank is equal to the number of nonzero linearly independent vectors.

4. The second and third rows are independent. However, the sum of the second and third rows are equal to the first: $(1, 0, 1) + (0, 1, 0) = (1, 1, 1)$. So this matrix has rank 2.

5. All three rows are independent. An easy way to check is to notice that this matrix is **upper triangular** , with nonzero entries along the diagonal.

The rank of a matrix continued

0/2 points (graded)

This question is meant to serve as an answer to the following: *If you sum two rank-1 matrices, do you get a rank-2 matrix? What about products? More generally, what rank is the sum of a rank-r_1 and a rank-r_2 matrix?*"

Let $\mathbf{A} = \begin{pmatrix} -1 & 1 \\ -3 & 3 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Observe that all four of these matrices are rank 1.

There are many ways to determine rank. Here is one useful fact that you could use for this problem:

> **"Every rank-1 matrix can be written as an outer product. Conversely, every outer product $\mathbf{u}\mathbf{v}^T$ is a rank-1 matrix."**

For example, $\mathbf{A} = \mathbf{u}\mathbf{v}^T$, $\mathbf{B} = \mathbf{v}\mathbf{v}^T$, $\mathbf{C} = \mathbf{w}\mathbf{w}^T$ and $\mathbf{D} = \mathbf{x}\mathbf{x}^T$, where

$$\mathbf{u} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Solution:

The choices are of two general types: sums of matrices, and products of matrices.

- $\mathbf{A} + \mathbf{A} = 2\mathbf{A}$, which has rank 1.
- $\mathbf{A} + \mathbf{B} = \mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{v}^T = (\mathbf{u} + \mathbf{v})\, \mathbf{v}^T$, which has rank 1.
- $\mathbf{A} + \mathbf{C} = \begin{pmatrix} -1 & 1 \\ -3 & 4 \end{pmatrix}$. This has two linearly independent rows, hence its rank is 2.

The last three choices $\mathbf{AB}$, $\mathbf{AC}$, $\mathbf{BD}$ cannot have rank 2 since they are products of rank-1 matrices.

- $\mathbf{AB} = \mathbf{u}\mathbf{v}^T\mathbf{v}\mathbf{v}^T = \mathbf{u}\langle\mathbf{v}, \mathbf{v}\rangle\mathbf{v}^T = \langle\mathbf{v}, \mathbf{v}\rangle\mathbf{u}\mathbf{v}^T$. Note that the inner product $\mathbf{v}^T\mathbf{v} = \langle\mathbf{v}, \mathbf{v}\rangle$ "floats" to the front because it is a scalar. This is an outer product of two vectors, which has rank 1.
- $\mathbf{AC} = \mathbf{u}\mathbf{v}^T\mathbf{w}\mathbf{w}^T = \langle\mathbf{v}, \mathbf{w}\rangle\mathbf{u}\mathbf{w}^T$, which again has rank 1.
- $\mathbf{BD} = \mathbf{v}\mathbf{v}^T\mathbf{x}\mathbf{x}^T = \langle\mathbf{v}, \mathbf{x}\rangle\mathbf{v}\mathbf{x}^T$. Notice that $\mathbf{v}$ is orthogonal to $\mathbf{x}$, so $\mathbf{BD} = 0\mathbf{v}\mathbf{x}^T$ is the zero matrix. Its rank is zero.

In general, the sum of two matrices can have a varying range of ranks, and they can be greater **or** less than the ranks of matrices that are being summed up. On the other hand, it is a general fact that if $\mathbf{A}$ and $\mathbf{B}$ are arbitrary (possibly rectangular) matrices, $\text{rank}\,(\mathbf{AB}) \leq \min\,(\text{rank}\,(\mathbf{A})\,, \text{rank}\,(\mathbf{B}))$. It is possible to use **determinants** to reason about rank. For choices such as $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 0 & 0 \\ -2 & 2 \end{pmatrix}$, the rank is obviously 1. Sometimes, it is easier if you know how to factor matrices – in this problem, we gave you the factorizations of rank-1 matrices into outer products of vectors. Other times, one may resort to using Gaussian Elimination – the rank of any upper triangular matrix is **at least** the number of non-zero entries along the diagonal.

Invertibility of a matrix

0 points possible (ungraded)

An $n \times n$ matrix $\mathbf{A}$ is invertible if and only if $\mathbf{A}$ has full rank, i.e. $\text{rank}\,(\mathbf{A}) = n$.

Which of the following matrices are invertible? Choose all that apply.

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{D} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Solution:

We saw in a previous exercise that the rank of $\mathbf{A}$ is 1. The rank of $\mathbf{B}$ is 2, since $(1,2)$ and $(2,1)$ are linearly independent, since e.g. by Gaussian Elimination one obtains the reduced upper triangular matrix $\begin{pmatrix} 1 & 2 \\ 0 & 3/2 \end{pmatrix}$. In general, an upper triangular matrix with nonzero entries along the diagonal has full rank. By the same reasoning, $\mathbf{C}$ also has full rank. Finally, $\mathbf{D}$ does not have full rank, since $(\text{row } 1) + (\text{row } 2) + (\text{row } 3) = \vec{0}$.

9. Eigenvalues, Eigenvectors and Determinants(Optional)

Bookmark this page

Eigenvalues and Eigenvectors of a matrix (Optional)

0 points possible (ungraded)

Let $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$\mathbf{Av} = \lambda_1 \mathbf{v}$, where $\lambda_1 =$

3

Answer: 3 .

$\mathbf{Aw} = \lambda_2 \mathbf{w}$, where $\lambda_2 =$

2

Answer: 2 .

Therefore, $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ_1, and $\mathbf{w}$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ_2.

Solution:

$$\mathbf{Av} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} \implies \lambda_1 = 3$$

$$\mathbf{Aw} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \implies \lambda_2 = 2$$

Geometric Interpretation of Eigenvalues and Eigenvectors (Optional)

0 points possible (ungraded)

Let $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Recall from the previous exercise that $\mathbf{v}$ and $\mathbf{w}$ are eigenvectors of $\mathbf{A}$.

Suppose $\mathbf{x} = \mathbf{v} + 2\mathbf{w} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Then $\mathbf{Ax} = s\mathbf{v} + t\mathbf{w}$, where:

$s =$ 3 ✔ **Answer:** 3

and

$t =$ 0.5 ✖ **Answer:** 4 .

In particular, s describes the amount that $\mathbf{A}$ stretches $\mathbf{x}$ in the direction of $\mathbf{v}$, and $\frac{t}{2}$ (note the "2" in front of $\mathbf{w}$ in $\mathbf{x}$) describes the amount that $\mathbf{A}$ stretches $\mathbf{x}$ in the direction of $\mathbf{w}$.

Solution:

We have

$$\begin{aligned} \mathbf{Ax} &= \mathbf{A}(\mathbf{v} + 2\mathbf{w}) \\ &= \mathbf{Av} + 2\mathbf{Aw} \\ &= (3\mathbf{v}) + 2(2\mathbf{w}) \\ &= 3\mathbf{v} + 4\mathbf{w}. \end{aligned}$$

From this, we get $s = 3, t = 4$.

Determinant and Eigenvalues (optional)

0 points possible (ungraded)

Recall that the **determinant** of a matrix indicates whether it is singular. For 2×2 matrices, it has the formula

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

but for larger matrices, the formula is more complicated.

What is the determinant of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix}$?

6

✔ **Answer:** 6

On the other hand, what is the product of the eigenvalues λ_1, λ_2 of $\mathbf{A}$? (We already computed this in the previous exercises.)

6

✔ **Answer:** 6

Solution:

Plugging into the formula directly gives $3 \cdot 2 - 0 \cdot \frac{1}{2} = 6$. On the other hand, the eigenvalues are $\lambda_1 = 3$, $\lambda_2 = 2$, so the product is 6. This is not a coincidence; for general $n \times n$ matrices, the **product of the eigenvalues is always equal to the determinant**.

Trace and Eigenvalues

0 points possible (ungraded)

Recall that the **trace** of a matrix is the sum of the diagonal entries.

What is the trace of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ \frac{1}{2} & 2 \end{pmatrix}$?

5

✔ **Answer:** 5

On the other hand, what is the sum of the eigenvalues λ_1, λ_2 of $\mathbf{A}$? (We already computed this in the previous exercises.)

5

✔ **Answer:** 5

Solution:

The diagonal sum is $3 + 2 = 5$. On the other hand, the eigenvalues are $\lambda_1 = 3$, $\lambda_2 = 2$, so the sum is 5. Just like the determinant, this is also not a coincidence. For general $n \times n$ matrices, the **sum of the eigenvalues is always equal to the trace of the matrix**.

Nullspace (Optional)

0 points possible (ungraded)

If a (nonzero) vector is in the nullspace of a square matrix $\mathbf{A}$, is it an eigenvector of $\mathbf{A}$?

yes

✔ **Answer:** yes

Which of the following are equivalent to the statement that 0 is an eigenvalue for a given square matrix $\mathbf{A}$? (Choose all that apply.)

- [ ] There exists a nonzero solution to $\mathbf{A}\mathbf{v} = \mathbf{0}$. ✔
- [x] $\det(\mathbf{A}) = 0$ ✔
- [ ] $\det(\mathbf{A}) \neq 0$
- [ ] $\mathrm{NS}(\mathbf{A}) = \mathbf{0}$
- [x] $\mathrm{NS}(\mathbf{A}) \neq \mathbf{0}$ ✔

✖

Solution:

- If a vector $\mathbf{v}$ is in the nullspace of $\mathbf{A}$, then $\mathbf{A}\mathbf{v} = \mathbf{0} = (0)\,\mathbf{v}$. So it is an eigenvector of $\mathbf{A}$ associated to the eigenvalue 0.
- If 0 is an eigenvalue for a matrix $\mathbf{A}$, then by definition, there exists a nonzero solution to $\mathbf{A}\mathbf{v} = \mathbf{0}$; that is, $\mathrm{NS}(\mathbf{A}) \neq \mathbf{0}$, and this only happens if and only if $\det(\mathbf{A}) = 0$.