Performance Analysis 2024-25 (IN4-341): basic tools

Caterina De Bacco

November 12, 2024

1 Probability

CONDITIONAL VS JOINT PROBABILITY

$$P(A,B) = P(A|B)\,P(B) \quad . \tag{1}$$

This is often used to compute conditional probabilities.

BAYES' RULE

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A)} \quad . \tag{2}$$

This is often used to compute conditional probabilities if you have the other conditional $P(A|B)$ and the prior $P(B)$.

EXPECTED VALUE

$$\mathbb{E}[X] = \sum_{x \in D} x\,P(X=x) \quad , \tag{3}$$

where D is the domain of the random variable X, i.e. the set of all possible values that it can take. This is valide for discrete variables.
For continuous variables you replace sums with integrals:

$$\mathbb{E}[X] = \int_{x \in D} x\,P(X=x)\,dx \quad . \tag{4}$$

CUMULATIVE DISTRIBUTION FUNCTION (CDF)

$$F_X(x) := P(X \leq x) \quad . \tag{5}$$

For continuous variables:

$$F_X(x) = \int_{-\infty}^{x} P(x)\,dx \quad ; \tag{6}$$

For discrete variables:

$$F_X(x) = \sum_{x_k \leq x} P(x_k) \quad . \tag{7}$$

This is often used to compute the probability density function $P(X)$, thanks to the relation:

$$P(X=x) = \frac{dF(x)}{dx} \quad . \tag{8}$$

For instance, if you have the following CDF:

$$F_X(x) = 1 - \mathrm{e}^{-\lambda x} \quad , \tag{9}$$

then you know that the $P(X)$ is the exponential distribution:

$$P(X=x) = \lambda\,\mathrm{e}^{-\lambda x} \quad . \tag{10}$$

Sometime is easier to characterize the $F_X(x)$.

COMPLEMENTARY CDF This is simply the other tail of the distribution:

$$\bar{F}_X(x) := P(X > x) = 1 - F_X(x) \quad . \tag{11}$$

In the continuous case:

$$\bar{F}_X(x) = \int_x^\infty P(x)\,dx \quad . \tag{12}$$

This is often useful to compute expected values, because of the relation:

$$\mathbb{E}[X] = \int_0^\infty \bar{F}_X(x)\,dx - \int_{-\infty}^0 F_X(x)\,dx \quad . \tag{13}$$

For positive variables this simplifies to:

$$\mathbb{E}[X] = \int_0^\infty \bar{F}_X(x)\,dx \quad ; \tag{14}$$

or, for discrete positive variables:

$$\mathbb{E}[X] = \sum_{n=0}^\infty \bar{F}_X(x_n) \quad , \tag{15}$$

because the second term in eq. (13) is 0.

GENERATING FUNCTION (pgf) For a *discrete* random variable X:

$$\rho_X(z) := \mathbb{E}\left[z^X\right] = \sum_x z^x\, Pr[X = x] \quad . \tag{16}$$

For a *continuous* random variable X:

$$\rho_X(z) := \mathbb{E}\left[e^{-zX}\right] = \int_{-\infty}^\infty e^{-zx} f_X(x)\,dx \quad . \tag{17}$$

Often one selects $z = e^t$ in the discrete case, because this allows to use the pgf to calculate moments of the distribution via:

$$\mathbb{E}[X^n] = \frac{d^n \rho_X(e^t)}{dt^n}\Big|_{t=0} \quad . \tag{18}$$

For instance, $\mathbb{E}[X] = \rho_X'(1)$ and $\mathbb{E}\left[X^2\right] = \rho_X''(1) + \rho_X'(1)$.

2 Linear algebra

EIGENVALUE EQUATION

$$Ax = \lambda x \quad , \tag{19}$$

where A is a matrix of dimension $N \times N$, x is a vector of dimension $N \times 1$ called (right) eigenvector, and $\lambda \in \mathbb{R}$ is a scalar number, called eigenvalue.
This is useful because many equations end up being of this type, e.g. computing the steady state of a stochastic process.
Typically one has access to A, and wants to find out what the eigenvectors and eigenvalues are.
If you use row vectors, instead of columns, you could also find the left eigenvector by transposing eq. (19) (and switching the order to xA):

$$A^T x^T = \lambda x^T \quad , \tag{20}$$

this also mean that the left eigenvector of A is the same as the transpose of a right eigenvector of A^T.

EIGENDECOMPOSITION

$$A = X D Y^T \quad , \tag{21}$$

where X, Y have as columns the right and left eigenvectors of A, and D is a diagonal matrix with the eigenvalues of A.
This is useful because you can rewrite expressions containing A in this decompose form, simplifying terms.

PERRON–FROBENIUS THEOREM There is more than one version. Here we focus on the case of irreducible non-negative matrices (which will be the main one in this class):
For an irreducible non-negative matrix A, the largest eigenvalue λ_{max} is positive, all other eigenvalues are $|\lambda| \leq \lambda_{max}$. The eigenvector of λ_{max} is a vector with all positive entries. This is also the only eigenvector with all positive entries, and the only one associated to λ_{max}.
For stochastic matrices, e.g. those representing probability distributions, we have that the normalization constraint $\sum_j P_{ij} = 1$ implies that the Perron–Frobenius theorem specializes to: $\lambda_{max} = 1$ *and its eigenvector is* $u = (1, \ldots, 1)$*, i.e. the vector with all entries equal to 1.*