Machine Learning for Engineering and Science Applications

Some relations for
Expectation and Covariance
(Slightly advanced)

Covariance of two vectors

Recall

$$\mathrm{Cov}[f(\mathrm{x}), g(\mathrm{y})] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)]\,)]$$

$$\mathrm{Cov}[\mathrm{x}, \mathrm{y}] = \overline{xy} - \overline{x}\,\overline{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

$$\mathrm{Cov}[\boldsymbol{x}, \boldsymbol{x}]_{i,j} = \mathrm{Cov}\left[\boldsymbol{x}_i, \boldsymbol{x}_j\right]$$

Let $\boldsymbol{\mu} \in \mathbb{R}^{\boldsymbol{n}}$ be the expectation of the random vector $\boldsymbol{x}$

Then,

$$\boxed{\mathrm{Cov}[x, x] = \mathbb{E}[(x - \mu)(x - \mu)^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T} \in \mathbb{R}^{m \times m}$$

We use the notation $\mathrm{Var}[x] = \mathrm{Cov}[x, x]$

Similarly, $\boxed{\mathrm{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \mathrm{Cov}[y, x]^T} \in \mathbb{R}^{m \times n}$

Ref : “Mathematics for Machine Learning” https://mml-book.com

Sums of random variables

- Recall that expectation is a linear operator

$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$

- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,

$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variances are a bit more involved.

$$\mathrm{Var}[x + y] = \mathrm{Var}[x] + \mathrm{Var}[y] + \mathrm{Cov}[x, y] + \mathrm{Cov}[y, x]$$

$$\mathrm{Var}[\alpha x] = \alpha \mathrm{Var}[x]$$

Note that, if x, y are independent then $\mathrm{Var}[x + y] = \mathrm{Var}[x] + \mathrm{Var}[y]$

What is $\mathrm{Var}[x - y]$?

Affine transformations

- Affine transform : Transformation of variables of the form $\boldsymbol{y} = A\boldsymbol{x} + b$
- Question: Given the mean and variance matrix of $\boldsymbol{x}$ find them for $\boldsymbol{y}$
- Then, since expectation is a linear operator

$$\mathbb{E}_y[\boldsymbol{y}] = \mathbb{E}_{\boldsymbol{x}}[A\boldsymbol{x} + \boldsymbol{b}] = A\mathbb{E}_x[\boldsymbol{x}] + \boldsymbol{b} = A\boldsymbol{\mu} + \boldsymbol{b}$$

$$\begin{aligned}
\mathbb{V}_y[y] &= \mathbb{V}_x[Ax + b] \\
&= \mathbb{V}_x[Ax] \\
&= \mathrm{Cov}[Ax, Ax] \\
&= \mathbb{E}[(Ax)(Ax)^T] - \mathbb{E}[Ax]\mathbb{E}[Ax]^T \\
&= \mathbb{E}[Axx^T A^T] - A\mathbb{E}[\mathrm{x}]\mathrm{E}[\mathrm{x}]^{\mathrm{T}}\mathrm{A}^{\mathrm{T}} \\
&= A(\mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T)A^T \\
&= A\mathrm{Cov}[x, x]A^T = A\Sigma A^T
\end{aligned}$$

Exercise : Prove similarly that $\mathrm{Cov}[x, y] = \Sigma A^T$