Correlation

Abhinav Mehta

April 26, 2020

1 Correlation matrix

Let us assume an n-dimensional stochastic process $\vec{X} = [X^1, X^2 \dots X^n]^T$ which are sampled at m-points $t_1, t_2, ..., t_m$. Then,

$$\vec{X} = \begin{bmatrix} X^1(t) \\ X^2(t) \\ \vdots \\ X^n(t) \end{bmatrix},$$

$$\vec{X}\vec{X}^T = \begin{bmatrix} X^1(t) \\ X^2(t) \\ \vdots \\ X^n(t) \end{bmatrix} \begin{bmatrix} X^1(t) & X^2(t) & \dots & X^n(t) \end{bmatrix}$$

$$= \begin{bmatrix} X^1(t)X^1(t) & X^1(t)X^2(t) & \dots & X^1(t)X^n(t) \\ X^2(t)X^1(t) & X^2(t)X^2(t) & \dots & X^2(t)X^n(t) \\ \vdots & \vdots & \ddots & \vdots \\ X^n(t)X^1(t) & X^n(t)X^2(t) & \dots & X^n(t)X^n(t) \end{bmatrix}$$

Now, we observe that,

$$cov(\vec{X}) = E(\vec{X}\vec{X}^T) - E(\vec{X})E(\vec{X}^T)$$

$$= \begin{bmatrix} E((X^1)^2) - E(X^1)^2 & E(X^1X^2) - E(X_1)E(X_2) & \dots & E(X^1X^n) - E(X_1)E(X_n) \\ E(X^2X^1) - E(X^2)E(X^1) & E((X^2)^2) - (E(X^2))^2 & \dots & E(X^2X^n) - E(X^2)E(X^n) \\ \vdots & \vdots & \ddots & \vdots \\ E(X^nX^1) - E(X^n)E(X^1) & E(X^nX^2) - E(X^n)E(X^2) & \dots & E((X^n)^2) - (E(X^n))^2 \end{bmatrix}$$

$$= \begin{bmatrix} var(X^1) & cov(X^1, X^2) & \dots & cov(X^1, X^n) \\ cov(X^2, X^1) & var(X^2) & \dots & cov(X^2, X^n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X^n, X^1) & cov(X^n, X^2) & \dots & var(X^n) \end{bmatrix}$$

We can trivially see that the covariance matrix is symmetric. We can also prove that the covariance matrix is positive semi-definite. Let us take a random variable $Y = \sum_i^n a_i X^i$ where $\vec{a} = [a_1, a_2, \dots a_n] \in \mathcal{R}^n$. Then,

$$
\begin{aligned}
var(Y) &= var(\sum_{i}^{n} a_i X^i) \\
&= cov(\sum_{i}^{n} a_i X^i, \sum_{i}^{n} a_i X^i) \\
&= \sum_{j}^{n}\sum_{i}^{n} a_i a_j cov(X_i, X_j) \qquad = [a_1, a_2, \dots a_n]^T cov(\vec{X})[a_1, a_2, \dots a_n] = \vec{a}^T cov(\vec{X})\vec{a}
\end{aligned}
$$

Since $var(Y)$ is always positive and $\vec{a}$ was chosen arbitrarily from $\mathcal{R}^n$. And,

$$var(Y) = \vec{a}^T cov(\vec{X})\vec{a} \geq 0$$

implies that the covariance matrix is positive semi-definite.

Covariance can be standardized to correlation by dividing with their corresponding standard deviations i.e.

$$corr(\vec{X}) = D^{-1} cov(\vec{X}) (D^{-1})^T, \text{ where } D = diag(cov(\vec{X}))$$

2 Cholesky decomposition

A matrix A is symmetric (i.e. $A = A^T$), positive-definite matrix ($x^T A x > 0, \forall x$) iff it is possible to write it as a product of unique lower triangular matrix L with positive diagonal values and its transpose i.e.

$$A = LL^T$$

where,

$$
\begin{aligned}
L_{j,j} &= \sqrt{A_{j,j} - \sum_{k=1}^{j-1} L_{j,k}^2}, \\
L_{i,j} &= \frac{1}{L_{j,j}}\left(A_{i,j} - \sum_{k=1}^{j-1} L_{i,k}L_{j,k}\right), \text{ where } i > j
\end{aligned}
$$

Given a covariance matrix C (correlation is $D^{-1}C(D^{-1})^T$), we would like to apply a transformation on $\vec{Y}$ where Y_i's are uncorrelated with 0 mean and unit variance, so that replicates the relationships in C. Since C is symmetric and positive-definite, we can easily decompose C as

$$C = LL^T$$

Now, we apply the transformation L on Y giving $Z = LY$. Computing the covariance for Z, we have,

$$
\begin{aligned}
corr(\vec{Z}) &= D^{-1} cov(\vec{Z}) (D^{-1})^T \\
&= D^{-1}[E(ZZ^T) - E(Z)E(Z^T))](D^{-1})^T \\
&= D^{-1}[E(LYY^TL^T) - E(LY)E(Y^TL^T)](D^{-1})^T \\
&= D^{-1}(LE(YY^T)L^T)(D^{-1})^T \\
&= D^{-1}(LIL^T)(D^{-1})^T \\
&= D^{-1}C(D^{-1})^T
\end{aligned}
$$