Multivariate normal distribution

Danny Nygård Hansen

14th October 2021

Definition 0.1: *Multivariate normal distribution*

An n-dimensional random vector X follows an n-variate normal distribution if its characteristic function is given by

$$\varphi_{\mathsf{X}}(t) = \exp\Bigl(\mathrm{i}\langle t, \mu\rangle - \frac{1}{2} t^\top \Sigma t\Bigr), \qquad t \in \mathbb{R}^n,$$

for some $\mu \in \mathbb{R}^n$ and an $n \times n$ positive semidefinite matrix Σ. In this case we write $\mathsf{X} \sim N_n(\mu, \Sigma)$.

Recall that a Σ is said to be positive semidefinite if $t^\top \Sigma t \geq 0$ for all $t \in \mathbb{R}^n$.

Proposition 0.2

Let $\mathsf{X} \sim N_n(\mu, \Sigma)$, $a \in \mathbb{R}^n$ and $B \in \mathrm{M}_{p\times n}(\mathbb{R})$. Then $a + B\mathsf{X} \sim N_n(a + B\mu, B\Sigma B^\top)$. In particular, $\langle t, \mathsf{X}\rangle \sim N(\langle t, \mu\rangle, t^\top \Sigma t)$ for all $t \in \mathbb{R}^n$.

Conversely, if X is an n-dimensional random variable, and $\langle t, \mathsf{X}\rangle$ is normally distributed for all $t \in \mathbb{R}^n$, then X is n-variate normally distributed. In fact, if there exist $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathrm{M}_n(\mathbb{R})$ such that $\langle t, \mathsf{X}\rangle \sim N(\langle t, \mu\rangle, t^\top \Sigma t)$ for all $t \in \mathbb{R}^n$, then $\mathsf{X} \sim N_n(\mu, \Sigma)$.

Proof. First notice that $B\Sigma B^\top$ is clearly positive semidefinite since Σ is. Furthermore,

$$\begin{aligned}
\varphi_{a+B\mathsf{X}}(s) &= \mathrm{e}^{\mathrm{i}\langle s,a\rangle} \exp\Bigl(\mathrm{i}\langle B^\top s, \mu\rangle - \frac{1}{2}(B^\top s)^\top \Sigma (B^\top s)\Bigr) \\
&= \exp\Bigl(\mathrm{i}\langle s, a + B\mu\rangle - \frac{1}{2} s^\top (B\Sigma B^\top) s\Bigr)
\end{aligned}$$

for all $s \in \mathbb{R}^n$ as desired.

For the converse direction, first assume that $\langle t, \mathsf{X}\rangle$ is normally distributed for all $t \in \mathbb{R}^n$. In particular, the coordinates $\mathsf{X}_1, \ldots, \mathsf{X}_n$ are normally distributed, and hence X has a well-defined mean vector μ and covariance matrix Σ.

Furthermore, we have $\mathbb{E}[\langle t, \mathsf{X}\rangle] = \langle t, \mu\rangle$ and $\mathrm{Var}[\langle t, \mathsf{X}\rangle] = t^\top \Sigma t$. (See Andersen.) It follows that

$$\varphi_{\mathsf{X}}(t) = \varphi_{\langle t, \mathsf{X}\rangle}(1) = \exp\Big(\mathrm{i}\langle t, \mu\rangle - \frac{1}{2} t^\top \Sigma t\Big)$$

for $t \in \mathbb{R}^n$. The final claim follows similarly. □

COROLLARY 0.3

If $\mathsf{X} \sim N_n(\mu, \Sigma)$, then μ is the mean vector and Σ the covariance matrix for X.

REMARK 0.4. Let $\mu \in \mathbb{R}^n$, and let $\Sigma \in \mathrm{M}_n(\mathbb{R})$ be positive semidefinite. We show that there exists a random variable X with distribution $N_n(\mu, \Sigma)$.

First let $\Sigma^{1/2} \in \mathrm{M}_n(\mathbb{R})$ be the (unique) positive semidefinite matrix such that $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$. We might construct $\Sigma^{1/2}$ as follows: Since Σ is symmetric there exists a diagonal matrix D and an orthogonal matrix Q such that $\Sigma = Q^\top D Q$. Since Σ is positive semidefinite, the entries in D are non-negative. Call the diagonal entries $\lambda_1, \dots, \lambda_n$. Let $D^{1/2}$ be the diagonal matrix whose entries along the diagonal are $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$, and let $\Sigma^{1/2} = Q^\top D^{1/2} Q$.

Now let $\mathsf{U}_1, \dots \mathsf{U}_n$ be i.i.d. $N(0,1)$-distributed variables, and let $\mathsf{U} = (\mathsf{U}_1, \dots, \mathsf{U}_n)$. For $t = (t_1, \dots, t_n) \in \mathbb{R}^n$, the characteristic function of U is then

$$\varphi_{\mathsf{U}}(t) = \prod_{j=1}^n \varphi_{\mathsf{U}_j}(t_j) = \prod_{j=1}^n \exp\Big(-\frac{1}{2} t_j^2\Big) = \exp\Big(-\frac{1}{2} t^\top I t\Big),$$

where I is the identity matrix. It follows that $\mathsf{U} \sim N_n(0, I)$. Letting $\mathsf{X} = \mu + \Sigma^{1/2}\mathsf{U}$, we find that $\mathsf{X} \sim N_n(\mu, \Sigma)$. ⌟