Multivariate CLT proof with characteristic functions

Kadin Zhang

21-721, Spring 2024

1 Preliminaries

Definition (Characteristic function): Let X be a random variable, then its *characteristic function* is

$$\phi_X(t) = \mathbf{E}\exp(itX).$$

If X is random vector in $\mathbb{R}^n$, define

$$\phi_X(t) = \mathbf{E}\exp(it^\top X).$$

Proposition (Characteristic function for Gaussian): The characteristic function of a random vector $X \sim N(\mu, M)$ is

$$\phi_X(t) = \exp(it^\top \mu - t^\top M t/2).$$

Proposition (Uniqueness): Let $\mu \in \mathbb{R}^n$ and Σ a positive semi-definite matrix. There is unique Gaussian distribution with mean μ and covariance Σ. Further, if $\det \Sigma > 0$ then the distribution has well defined density

$$p(t) = \frac{1}{(2\pi)^{n/2} \det \Sigma^{1/2}} \exp\left(-\frac{1}{2}(t-\mu)^\top \Sigma^{-1}(t-\mu)\right).$$

Theorem (Levy continuity theorem): Let X_n have associated characteristic functions ϕ_n.

(a) Suppose X_n converges in distribution to some RV X. Then $\phi_n \to \phi_X$ pointwise.

(b) Suppose ϕ_n converges pointwise to a function ϕ which is continuous at 0. Then ϕ is characteristic function of some RV X and $X_n \xrightarrow{\mathrm{d}} X$.

2 Central limit theorem

Theorem (Multivariate CLT): Let X_n be sequence of i.i.d. random vectors in $\mathbb{R}^n$ with mean μ and covariance Σ. Then the sequence of random vectors

$$\frac{X_1 + \cdots + X_m - m\mu}{\sqrt{m}}$$

converges to a centered multivariate normal with covariance Σ.

Proof. We prove the theorem for $\mu = 0$, $\Sigma = I$, as the general result holds by linear transformation. First consider the one dimensional case. Let $Z_m = \frac{X_1 + \cdots + X_m}{\sqrt{m}}$. Then, using first the product rule then Taylor's theorem around 0,

$$\begin{aligned}\phi_{Z_m}(t) &= \phi_X(t/\sqrt{m})^m \\ &= \left(1 + \frac{t}{\sqrt{m}}\phi'_X(0) + \frac{t^2}{2m}\phi''_X(0) + o(1/m)\right)^m \\ &= \left(1 - \frac{t^2}{2m} + o(1/m)\right)^m && (\phi_X \text{ gives moments}) \\ &\to \exp(-t^2/2).\end{aligned}$$

It follows that $\phi_{Z_m}(t)$ converges pointwise to the characteristic function of a standard Gaussian. Thus Z_m converges in distribution to Z, where $Z \sim N(0,1)$.

In the general case, fix $t \in \mathbb{R}^n$ and define $Y_m = t^\top X_m$. So Y_m are i.i.d. random variables with mean 0 and variance

$$\mathbf{E}(t^\top X)^2 = t^\top \mathbf{E}(X t^\top X) = t^\top \mathbf{E}(X X^\top t) = t^\top t.$$

Apply the one dimensional CLT to Y_m to get

$$\frac{Y_1 + \cdots + Y_m}{\sqrt{m}} \xrightarrow{\mathrm{d}} Z,$$

where $Z \sim N(0, t^\top t)$. Then, by Levy continuity theorem, for $s \in \mathbb{R}$, letting $W_m = \frac{Y_1 + \cdots + Y_m}{\sqrt{m}}$,

$$\phi_{W_m}(s) = \phi_{Y_m}(s/\sqrt{m})^m \to \phi_Z(s).$$

Note that

$$\phi_{X_m}(t) = \exp(it^\top X_m) = \exp(iY_m) = \phi_{Y_m}(1).$$

Therefore, setting $s = 1$, and letting $Z_m = \frac{X_1 + \cdots + X_m}{\sqrt{m}}$,

$$\phi_{Z_m}(t) = \phi_{X_m}(t/\sqrt{m})^m \to \exp(-t^\top t/2).$$

So again using Levy continuity theorem, $Z_m \xrightarrow{\mathrm{d}} Z$, where $Z \sim N(0, I)$.

□