Natural Language Processing

1 KL divergence of Gaussians

1 Dimension:

$$D_{\mathrm{KL}}(P\|Q) = \frac{1}{2}\left[\log\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} + \frac{(\mu_1-\mu_2)^2}{\sigma_2^2} - 1\right]$$

D Dmension:

$$D_{\mathrm{KL}}(P\|Q) = \frac{1}{2}\left[\log\frac{|\Sigma_2|}{|\Sigma_1|} + \mathrm{tr}\left\{\Sigma_2^{-1}\Sigma_1\right\} + (\vec{\mu}_1-\vec{\mu}_2)^{\mathrm{T}}\Sigma_2^{-1}(\vec{\mu}_1-\vec{\mu}_2) - d\right]$$

Reference