

1 Objectives

Calculate mutual information using numerical integration for a complex, linearized signal model $\vec{z}(\mu, \vec{k})$.

$$\vec{z}(\mu, \vec{k}) = \int_{\Omega} M(\vec{x}) e^{-s(\mu, \vec{x})} e^{-2\pi i \vec{k} \cdot \vec{x}} d\vec{x} + \nu = \mathcal{G}(\mu, \vec{k}) + \nu \quad \nu \sim \mathcal{N}(\vec{0}, \Sigma_{\nu}) \quad \Sigma_{\nu} = \begin{bmatrix} \sigma_{\nu, r}^2 & 0 \\ 0 & \sigma_{\nu, i}^2 \end{bmatrix} \quad (1)$$

Assume normal distribution for the model parameter, optical attenuation coefficient μ .

$$\mu \sim \mathcal{N}(\vec{\mu}_{\mu}, \Sigma_{\mu}) \quad \vec{\mu}_{\mu} = \begin{bmatrix} \mu_{\mu, r} \\ 0 \end{bmatrix} \quad \Sigma_{\mu} = \begin{bmatrix} \sigma_{\mu, r}^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (2)$$

2 Derivations

2.1 Multivariate Normal Distribution Representation

This is the part that is confusing. With the change of variables, $p(\mathbf{z}|\mu)$ goes from a normal distribution in two-dimensional vector variable \mathbf{z} to a normal distribution in one-dimensional scalar variable μ with two-dimensional mean and covariance matrix.

$$p(\mathbf{z}|\mu) = \mathcal{N}_{\mathbf{z}}(\mu_{\nu}, \Sigma_{\nu}) = \mathcal{N}_{\mu}(\mu_{\nu, 2}, \Sigma_{\nu, 2}) \quad (3)$$

2.2 Complex Normal Distribution Representation

Given a k-dimensional complex random variable \mathbf{Z} :

$$\mathbf{Z} = \mathbf{X} + i\mathbf{Y} \quad \mathbf{X}, \mathbf{Y} \in \mathbb{R}^k \quad (4)$$

If \mathbf{Z} follows a complex normal distribution $\mathbf{Z} \sim \mathcal{CN}(\mu, \mathbf{\Gamma}, \mathbf{C})$, the distribution is defined by three parameters, the mean μ , the covariance matrix $\mathbf{\Gamma}$, and the similarity matrix \mathbf{C} . These are defined as follows, where the bar represents the complex conjugate.

$$\mu = \mathbb{E}[\mathbf{Z}] \quad \mathbf{\Gamma} = \mathbb{E}[(\mathbf{Z} - \mu)(\bar{\mathbf{Z}} - \bar{\mu})^T] \quad \mathbf{C} = \mathbb{E}[(\mathbf{Z} - \mu)(\mathbf{Z} - \mu)] \quad (5)$$

Definition of probability distribution function for complex normal distribution:

$$p(\mathbf{z}) = \frac{1}{\pi^k \sqrt{\det(\mathbf{\Gamma}) \det(\bar{\mathbf{\Gamma}} - \bar{\mathbf{C}}^T \mathbf{\Gamma}^{-1} \mathbf{C})}} \exp \left\{ -\frac{1}{2} \begin{pmatrix} (\bar{\mathbf{Z}} - \bar{\mu})^T & (\mathbf{Z} - \mu)^T \end{pmatrix} \begin{pmatrix} \mathbf{\Gamma} & \mathbf{C} \\ \bar{\mathbf{C}}^T & \bar{\mathbf{\Gamma}} \end{pmatrix} \begin{pmatrix} \mathbf{Z} - \mu \\ \bar{\mathbf{Z}} - \bar{\mu} \end{pmatrix} \right\} \quad (6)$$

Covariance matrix in terms of variances and covariances of real vector \mathbf{X} and imaginary vector \mathbf{Y} :

$$\mathbf{\Gamma} = \mathbf{V}_{\mathbf{xx}} + \mathbf{V}_{\mathbf{yy}} + i(\mathbf{V}_{\mathbf{yx}} - \mathbf{V}_{\mathbf{xy}}) \quad (7)$$

Similarity matrix in terms of variances and covariances of real vector \mathbf{X} and imaginary vector \mathbf{Y} :

$$\mathbf{C} = \mathbf{V}_{\mathbf{xx}} - \mathbf{V}_{\mathbf{yy}} + i(\mathbf{V}_{\mathbf{yx}} + \mathbf{V}_{\mathbf{xy}}) \quad (8)$$

$$\mathbf{R} = \bar{\mathbf{C}}^T \mathbf{\Gamma}^{-1} \quad (9)$$

$$\mathbf{P} = \bar{\mathbf{\Gamma}} - \mathbf{R}\mathbf{C} = \bar{\mathbf{\Gamma}} - \bar{\mathbf{C}}^T \mathbf{\Gamma}^{-1} \mathbf{C} \quad (10)$$

The conditional distribution $p(\mathbf{z}|\mu)$ represented as a two-dimensional real normal distribution:

$$\mathbf{z}|\mu \sim \mathcal{N}(\mathcal{G}(\mu, \mathbf{k}), \Sigma_\nu) \quad \boldsymbol{\mu}_\nu = \begin{bmatrix} \text{Re}(\mathcal{G}(\mu, \mathbf{k})) \\ \text{Im}(\mathcal{G}(\mu, \mathbf{k})) \end{bmatrix} = \begin{bmatrix} a_r(\mu - \hat{\mu}) + b_r \\ a_i(\mu - \hat{\mu}) + b_i \end{bmatrix} = \begin{bmatrix} \mu_{\nu,r} \\ \mu_{\nu,i} \end{bmatrix} \quad \Sigma_\nu = \begin{bmatrix} \sigma_{\nu,r}^2 & 0 \\ 0 & \sigma_{\nu,r}^2 \end{bmatrix} = \begin{bmatrix} \sigma_\nu^2 & 0 \\ 0 & \sigma_\nu^2 \end{bmatrix} \quad (11)$$

The model parameter distribution $p(\mu)$ represented as a two-dimensional real normal distribution:

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu}_\mu, \Sigma_\mu) \quad \boldsymbol{\mu}_\mu = \begin{bmatrix} \mu_\mu \\ 0 \end{bmatrix} \quad \Sigma_\mu = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (12)$$

The conditional distribution $p(\mathbf{z}|\mu)$ represented as a one-dimensional complex normal distribution:

$$\begin{aligned} \mathbf{z}|\mu &\sim \mathcal{CN}(\mu_\nu, \Gamma_\nu, C_\nu) \\ \mu_\nu &= \mu_{\nu,r} + i\mu_{\nu,i} \\ \Gamma_\nu &= \sigma_\nu^2 + i\sigma_\nu^2 \\ C_\nu &= 0 \\ R_\nu &= 0 \\ P_\nu &= \sigma_n u^2 - i\sigma_\nu^2 \end{aligned} \quad (13)$$

The PDF is

$$p(\mathbf{z}|\mu) = \frac{1}{\pi^2 \sqrt{\det(\Gamma_\nu) \det(\bar{\Gamma}_\nu - \bar{0}^T \Gamma_\nu^{-1} 0)}} \exp \left\{ -\frac{1}{2} \begin{pmatrix} (\bar{\mathbf{z}} - \bar{\mu}_\nu)^T & (\mathbf{z} - \mu_\nu)^T \end{pmatrix} \begin{pmatrix} \Gamma_\nu & 0 \\ \bar{0}^T & \bar{\Gamma}_\nu \end{pmatrix} \begin{pmatrix} \mathbf{z} - \mu_\nu \\ \bar{\mathbf{z}} - \bar{\mu}_\nu \end{pmatrix} \right\} \quad (14)$$

The model parameter distribution $p(\mu)$ represented as a one-dimensional complex normal distribution:

$$\begin{aligned} \boldsymbol{\mu} &\sim \mathcal{CN}(\mu_\mu, \Gamma_\mu, C_\mu) \\ \mu_\mu &= \mu_\mu \\ \Gamma_\mu &= \sigma_\mu^2 \\ C_\mu &= \sigma_\mu^2 \\ R_\mu &= 1 \\ P_\mu &= 0 \end{aligned} \quad (15)$$

The PDF is

$$p(\mu) = \frac{1}{\pi^2 \sqrt{\det((\sigma_\mu^2)) \det((\bar{\sigma}_\mu^2) - (\bar{\sigma}_\mu^2)^T (\sigma_\mu^2)^{-1} (\sigma_\mu^2))}} \exp \left\{ -\frac{1}{2} \begin{pmatrix} (\bar{\boldsymbol{\mu}} - \bar{\mu}_\mu)^T & (\boldsymbol{\mu} - \mu_\mu)^T \end{pmatrix} \begin{pmatrix} (\sigma_\mu^2) & (\sigma_\mu^2) \\ (\bar{\sigma}_\mu^2)^T & (\bar{\sigma}_\mu^2) \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu} - \mu_\mu \\ \bar{\boldsymbol{\mu}} - \bar{\mu}_\mu \end{pmatrix} \right\} \quad (16)$$

3 Directions

$\boldsymbol{\mu}$ cannot be represented as a complex normal distribution. Because only its real part is distributed and its imaginary part is zero, the denominator blows up in the complex normal PDF. Representing $p(\mathbf{z}|\mu)$ as a complex normal distribution and $p(\boldsymbol{\mu})$ as a one-dimensional real normal distribution has no advantages over the other approaches tried so far.

It seems that the most straightforward way to solve this problem is through numerical integration. This will require nested Gauss-Legendre quadrature and Gauss-Hermite quadrature. It may also be wise to use a physically possible uniform distribution for $p(\mu)$ if an analytical solution is out of reach.

$$I(\mu; z) = H(z) - H(z|\mu) = \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\Sigma_z| \right) - \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\Sigma_\nu| \right) \quad (17)$$

$$\Sigma_z = \begin{bmatrix} \text{E} \left[(z_r - \bar{z}_r)^2 \right] & \text{E} \left[(z_r - \bar{z}_r)(z_i - \bar{z}_i) \right] \\ \text{E} \left[(z_i - \bar{z}_i)(z_r - \bar{z}_r) \right] & \text{E} \left[(z_i - \bar{z}_i)^2 \right] \end{bmatrix} \quad (18)$$

$E[z_r]$, $E[z_i]$, $E[(z_r - \bar{z}_r)^2]$, $E[(z_i - \bar{z}_i)^2]$, and $E[(z_r - \bar{z}_r)(z_i - \bar{z}_i)]$ all need to be calculated numerically. As one example of this process, take $E[z_r]$:

$$E[z_r] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r p(\mathbf{z}) dz_r dz_i = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r \left[\int_{-\infty}^{\infty} p(\mathbf{z}|\mu) p(\mu) d\mu \right] dz_r dz_i \quad (19)$$

$$E[z_r] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r \left[\int_{-\infty}^{\infty} \mathcal{N}_{z_r}(a_r(\mu - \hat{\mu}) + b_r, \sigma_r^2) \mathcal{N}_{z_i}(a_i(\mu - \hat{\mu}) + b_i, \sigma_i^2) \mathcal{N}_{\mu}(\mu, \sigma_{\mu}^2) d\mu \right] dz_r dz_i \quad (20)$$

OR

$$E[z_r] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r \left[\int_{-\infty}^{\infty} \mathcal{N}_{z_r}(a_r(\mu - \hat{\mu}) + b_r, \sigma_r^2) \mathcal{N}_{z_i}(a_i(\mu - \hat{\mu}) + b_i, \sigma_i^2) \mathcal{U}_{\mu}(a_{\mu}, b_{\mu}) d\mu \right] dz_r dz_i \quad (21)$$

Gaussian quadrature:

$$E[z_r] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r \left[\int_{-\infty}^{\infty} p(\mathbf{z}|\mu) p(\mu) d\mu \right] dz_r dz_i \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_r [\sum_{i=1}^n w_i p(\mathbf{z}|\mu_i) p(\mu_i)] dz_r dz_i \quad (22)$$