

Gauss-Hermite Quadrature for Mutual Information

Drew Mitchell

May 18, 2016

1 Probability Distribution Definitions

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

$$\begin{aligned}\mathbf{z}(\boldsymbol{\mu}, \mathbf{k}) &= \int_{\Omega} M(\mathbf{x}) e^{-s(\boldsymbol{\mu}, \mathbf{x})} e^{-2\pi i \mathbf{k} \cdot \mathbf{x}} d\mathbf{x} + \boldsymbol{\nu} \\ &= \mathcal{G}(\boldsymbol{\mu}, \mathbf{k}) + \boldsymbol{\nu} \\ s(\boldsymbol{\mu}, \mathbf{x}) &= \frac{T_E}{T_2^*(\mathbf{x})} + i [2\pi \gamma \alpha B_0 T_E \Delta u(\boldsymbol{\mu}, \mathbf{x}) + T_E \Delta \omega_0(\mathbf{x})] \\ \boldsymbol{\nu} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\nu}), \quad \boldsymbol{\Sigma}_{\nu} = \begin{bmatrix} \sigma_{\nu, r}^2 & 0 \\ 0 & \sigma_{\nu, i}^2 \end{bmatrix}\end{aligned}\tag{1}$$

Therefore, the probability distribution for $p(\mathbf{z} \mid \boldsymbol{\mu})$ is

$$p(\mathbf{z} \mid \boldsymbol{\mu}) = \mathcal{N}(\mathcal{G}(\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\nu})\tag{2}$$

The tissue properties can be described by the following piecewise functions.

$$\begin{aligned}\boldsymbol{\mu}(\mathbf{x}) &= \sum_{n=1}^N \boldsymbol{\mu}_n U(\mathbf{x} - \Omega_n) \\ \bigcup_{n=1}^N \Omega_n &= \Omega, \quad \Omega_n \cap \Omega_m = \emptyset \\ U(\mathbf{x} - \Omega_n) &= \begin{cases} 1, & x \in \Omega_n \\ 0, & \text{otherwise} \end{cases}\end{aligned}\tag{3}$$

Assume normal distribution for the model parameter, optical attenuation coefficient $\boldsymbol{\mu}$.

$$p(\boldsymbol{\mu}) = \mathcal{N}(\mathbf{m}_{\mu}, \boldsymbol{\Sigma}_{\mu})\tag{4}$$

$$\mathbf{m}_\mu = \begin{bmatrix} m_{\mu_1} \\ \vdots \\ m_{\mu_N} \end{bmatrix} \quad (5a)$$

$$\mathbf{\Sigma}_\mu = \begin{bmatrix} \sigma_{\mu_1} & 0 & \cdots & 0 & 0 \\ 0 & \sigma_{\mu_2} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \sigma_{\mu_{N-1}} & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{\mu_N} \end{bmatrix} \quad (5b)$$

Also of note, because the real and imaginary components of \mathbf{z} are assumed independent, the covariance matrix $\mathbf{\Sigma}_z$ is diagonal, and the following simplification results.

$$p(\mathbf{z} | \boldsymbol{\mu}) = \mathcal{N}_{\mathbf{z}}(\mathcal{G}(\boldsymbol{\mu}), \mathbf{\Sigma}_\nu) = \mathcal{N}_{z_r}(\mathcal{G}_r(\boldsymbol{\mu}), \mathbf{\Sigma}_\nu) \mathcal{N}_{z_i}(\mathcal{G}_i(\boldsymbol{\mu}), \mathbf{\Sigma}_\nu) = p(z_r | \boldsymbol{\mu}) p(z_i | \boldsymbol{\mu}) \quad (6)$$

Similarly, the various tissue types are independent, and the covariance matrix $\mathbf{\Sigma}_\mu$ is also diagonal.

$$p(\boldsymbol{\mu}) = \mathcal{N}_{\boldsymbol{\mu}}(\mathbf{m}_\mu, \mathbf{\Sigma}_\mu) = \prod_{i=1}^N \mathcal{N}_{\mu_i}(m_{\mu_i}, \sigma_{\mu_i}) = \prod_{i=1}^N p(\mu_i) \quad (7)$$

2 Problem Statement

The most approachable way to solve mutual information is to begin with the difference of entropies definition.

$$\begin{aligned} I(\boldsymbol{\mu}; \mathbf{z}) &= H(\mathbf{z}) - H(\mathbf{z} | \boldsymbol{\mu}) \\ &= \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\mathbf{\Sigma}_z| \right) - \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\mathbf{\Sigma}_\nu| \right) \end{aligned} \quad (8)$$

Assuming $\sigma_{\nu,r}^2 = \sigma_{\nu,i}^2$, then

$$|\mathbf{\Sigma}_\nu| = \sigma_\nu^2 \sigma_\nu^2 - 0 = \sigma_\nu^4 \quad (9)$$

Calculation of $\mathbf{\Sigma}_z$ is less straightforward.

$$\begin{aligned} \mathbf{\Sigma}_z &= \begin{bmatrix} \sigma_{rr} & \sigma_{ri} \\ \sigma_{ir} & \sigma_{ii} \end{bmatrix} \\ \sigma_{rr} &= \mathbb{E} \left[(z_r - \mathbb{E}[z_r])^2 \right] \\ \sigma_{ri} &= \mathbb{E} \left[(z_r - \mathbb{E}[z_r]) (z_i - \mathbb{E}[z_i]) \right] \\ \sigma_{ir} &= \mathbb{E} \left[(z_i - \mathbb{E}[z_i]) (z_r - \mathbb{E}[z_r]) \right] \\ \sigma_{ii} &= \mathbb{E} \left[(z_i - \mathbb{E}[z_i])^2 \right] \end{aligned} \quad (10)$$

The values of σ_{rr} , σ_{ri} , σ_{ir} , and σ_{ii} all need to be calculated numerically.

Alternatively, $\int p(z | \mu)p(\mu)$ can probably be calculated analytically using [Bromiley 2003]. The product of three normal distributions is a scaled normal distribution, so after converting N_{z_i} and N_{z_r} to distributions in terms of μ for the integration, the normal distribution integrates to 1, and the scaling factor comes out as a function of z_r and z_i , $S_{fgh}(z_r, z_i)$, for the remaining two integrals. This leaves potentially very difficult numerical integrations at this step.

3 Gauss-Hermite Quadrature

Gaussian quadrature:

$$\int_{-\infty}^{\infty} \exp^{-x^2} f(x) dx \approx \sum_{i=1}^N \omega_i f(x_i) \quad (11)$$

$$\omega_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(x_i)]^2}$$

where n is the number of sample points used, $H_n(x)$ is the physicists' Hermite polynomial, x_i are the roots of the Hermite polynomial, and ω_i are the associated Gauss-Hermite weights.

Substitution for normal distributions using Gauss-Hermite quadrature:

$$\mathbb{E}[h(y)] = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) h(y) dy \quad (12)$$

h is some function of y , and random variable Y is normally distributed.

$$x = \frac{y - \mu}{\sqrt{2}\sigma} \Leftrightarrow y = \sqrt{2}\sigma x + \mu \quad (13)$$

$$\begin{aligned} \mathbb{E}[h(y)] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \exp(-x^2) h(\sqrt{2}\sigma x + \mu) dx \\ &\approx \frac{q}{\sqrt{\pi}} \sum_{i=1}^N \omega_i h(\sqrt{2}\sigma x_i + \mu) \end{aligned} \quad (14)$$

4 Mutual Information Calculation

$$I(\mu; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z} | \mu) = \frac{1}{2} \ln\left((2\pi e)^2 \cdot |\Sigma_z|\right) - \frac{1}{2} \ln\left((2\pi e)^2 \cdot |\Sigma_\nu|\right) \quad (15)$$

$$|\Sigma_\nu| = \sigma_\nu^2 \sigma_\nu^2 - 0 = \sigma_\nu^4 \quad (16)$$

$$\mathbf{\Sigma}_z = \begin{bmatrix} \sigma_{rr} & \sigma_{ri} \\ \sigma_{ir} & \sigma_{ii} \end{bmatrix} \quad (17)$$

$$\begin{aligned} \sigma_{rr} &= \mathbb{E} \left[(z_r - \mathbb{E}[z_r])^2 \right] \\ \sigma_{ri} &= \mathbb{E} \left[(z_r - \mathbb{E}[z_r]) (z_i - \mathbb{E}[z_i]) \right] \\ \sigma_{ir} &= \mathbb{E} \left[(z_i - \mathbb{E}[z_i]) (z_r - \mathbb{E}[z_r]) \right] \\ \sigma_{ii} &= \mathbb{E} \left[(z_i - \mathbb{E}[z_i])^2 \right] \end{aligned}$$

$$p(\mathbf{z}) = \int p(\mathbf{z} | \boldsymbol{\mu}) p(\boldsymbol{\mu}) d\boldsymbol{\mu} \quad (18)$$

$$p(\mathbf{z} | \boldsymbol{\mu}) = \mathcal{N}_z(\mathcal{G}(\boldsymbol{\mu}), \mathbf{\Sigma}_\nu) = \mathcal{N}_{z_r}(\mathcal{G}_r(\boldsymbol{\mu}), \sigma_\nu) \mathcal{N}_{z_i}(\mathcal{G}_i(\boldsymbol{\mu}), \sigma_\nu) = p(z_r | \boldsymbol{\mu}) p(z_i | \boldsymbol{\mu}) \quad (19)$$

$$p(\boldsymbol{\mu}) = \mathcal{N}_\mu(\mathbf{m}_\mu, \mathbf{\Sigma}_\mu) = \prod_{i=1}^N \mathcal{N}_{\mu_i}(m_{\mu_i}, \sigma_{\mu_i}) = \prod_{i=1}^N p(\mu_i) \quad (20)$$

The signal measurement probability distribution $p(\mathbf{z})$ can be approximated using Gauss-Hermite quadrature. The calculation is performed separately for the real and imaginary components, $p(z_r)$ and $p(z_i)$, because z_r and z_i are assumed independent.

$$\begin{aligned} p(z_r) &= \int_{-\infty}^{\infty} p(z_r | \boldsymbol{\mu}) \prod_{n=1}^N p(\mu_n) d\boldsymbol{\mu} = \int_{-\infty}^{\infty} p(z_r | \boldsymbol{\mu}) \prod_{n=1}^N \frac{1}{\sqrt{2\pi}\sigma_{\mu_n}} \exp\left(-\frac{(\mu_n - m_{\mu_n})^2}{2\sigma_{\mu_n}^2}\right) d\boldsymbol{\mu} \\ &= \int_{-\infty}^{\infty} p(z_r | \boldsymbol{\mu}) \prod_{n=1}^N \frac{1}{\sqrt{\pi}} \exp(-x_n^2) d\mathbf{x} \approx \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p p(z_r | \boldsymbol{\mu}_p) \end{aligned} \quad (21)$$

Q=number of quadrature points; N=number of tissue types

$$\begin{aligned} p(z_r | \boldsymbol{\mu}) &= p(z_r | \mu_1, \dots, \mu_N) \\ p(z_r | \boldsymbol{\mu}_p) &= p(z_r | \mu_{1,q}, \dots, \mu_{N,q}) \\ d\boldsymbol{\mu} &= d\mu_1 \dots d\mu_N \\ \mu_{i,q} &= \sqrt{2}\sigma_{\mu_i} x_{i,q} + m_{\mu_i} \\ d\mu_{i,q} &= \sqrt{2}\sigma_{\mu_i} dx_{i,q} \end{aligned}$$

An identical calculation for the imaginary component results in

$$\begin{aligned} p(z_i) &\approx \pi^{-N/2} \sum_{i=1}^N \sum_{q=1}^Q \omega_{i,q} p(z_i | \boldsymbol{\mu}_p) \\ p(z_i | \boldsymbol{\mu}_p) &= p(z_i | \mu_{1,q}, \dots, \mu_{N,q}) \\ \mu_{i,q} &= \sqrt{2}\sigma_{\mu_i} x_{i,q} + m_{\mu_i} \end{aligned} \quad (22)$$

Calculate expectation values for real and imaginary components by definition and approximating $p(\mathbf{z})$ with Gauss-Hermite quadrature as shown above.

$$\begin{aligned}
\mathbb{E}[z_r] &= \int_{-\infty}^{\infty} z_r p(z_r) dz_r \approx \int_{-\infty}^{\infty} z_r \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p p(z_r | \boldsymbol{\mu}_p) dz_r \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \int_{-\infty}^{\infty} z_r p(z_r | \boldsymbol{\mu}_p) dz_r = \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathbb{E}[z_r | \boldsymbol{\mu}_p] = \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r(\boldsymbol{\mu}_p)
\end{aligned} \tag{23}$$

Similarly for the imaginary component,

$$\mathbb{E}[z_i] \approx \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_i(\boldsymbol{\mu}_p) \tag{24}$$

The variance can be calculated as follows:

$$\begin{aligned}
\mathbb{E}[(z_r - \mathbb{E}[z_r])^2] &= \mathbb{E}[z_r^2] - (\mathbb{E}[z_r])^2 = \int_{-\infty}^{\infty} z_r^2 p(z_r) dz_r - (\mathbb{E}[z_r])^2 \\
&\approx \int_{-\infty}^{\infty} z_r^2 \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p p(z_r | \boldsymbol{\mu}_p) dz_r - (\mathbb{E}[z_r])^2 \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \int_{-\infty}^{\infty} z_r^2 p(z_r | \boldsymbol{\mu}_p) dz_r - (\mathbb{E}[z_r])^2 = \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \\
&\quad \cdot \int_{-\infty}^{\infty} z_r^2 \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(-\frac{(z_r - \mathcal{G}_r(\boldsymbol{\mu}_p))^2}{2\sigma_\nu^2}\right) dz_r - (\mathbb{E}[z_r])^2 \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \int_{-\infty}^{\infty} (x^2 \sigma_\nu^2 + 2x\sigma_\nu \mathcal{G}_r(\boldsymbol{\mu}_p) + \mathcal{G}_r^2(\boldsymbol{\mu}_p)) \\
&\quad \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx - (\mathbb{E}[z_r])^2 \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p (\sigma_\nu^2 + \mathcal{G}_r^2(\boldsymbol{\mu}_p)) - (\mathbb{E}[z_r])^2 \\
&= \pi^{-N/2} \sigma_\nu^2 + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_r])^2
\end{aligned} \tag{25}$$

Again, an identical calculation for the imaginary component yields

$$\mathbb{E} \left[(z_i - \mathbb{E}[z_i])^2 \right] = \pi^{-N/2} \sigma_\nu^2 + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_i^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_i])^2 \quad (26)$$

The off-diagonal elements of Σ_z are equal.

$$\mathbb{E} [(z_r - \mathbb{E}[z_r]) (z_i - \mathbb{E}[z_i])] = \mathbb{E} [(z_i - \mathbb{E}[z_i]) (z_r - \mathbb{E}[z_r])] \quad (27)$$

$$\begin{aligned}
\mathbb{E}[(z_r - \mathbb{E}[z_r])(z_i - \mathbb{E}[z_i])] &= \int_{-\infty}^{\infty} (z_r - \mathbb{E}[z_r])(z_i - \mathbb{E}[z_i]) p(\mathbf{z}) d\mathbf{z} \\
&= \int_{-\infty}^{\infty} (z_r - \mathbb{E}[z_r])(z_i - \mathbb{E}[z_i]) \int_{-\infty}^{\infty} p(\mathbf{z} | \boldsymbol{\mu}) p(\boldsymbol{\mu}) d\boldsymbol{\mu} d\mathbf{z} \\
&\approx \int_{-\infty}^{\infty} (z_r - \mathbb{E}[z_r])(z_i - \mathbb{E}[z_i]) \cdot \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p p(\mathbf{z} | \boldsymbol{\mu}_p) d\mathbf{z} \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \int_{-\infty}^{\infty} (z_r - \mathbb{E}[z_r]) p(z_r | \boldsymbol{\mu}_p) dz_r \\
&\quad \cdot \int_{-\infty}^{\infty} (z_i - \mathbb{E}[z_i]) p(z_i | \boldsymbol{\mu}_p) dz_i \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \\
&\quad \cdot \int_{-\infty}^{\infty} (z_r - \mathbb{E}[z_r]) \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(-\frac{(z_r - \mathcal{G}_r(\boldsymbol{\mu}_p))^2}{2\sigma_\nu^2}\right) dz_r \\
&\quad \cdot \int_{-\infty}^{\infty} (z_i - \mathbb{E}[z_i]) \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(-\frac{(z_i - \mathcal{G}_i(\boldsymbol{\mu}_p))^2}{2\sigma_\nu^2}\right) dz_i \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \\
&\quad \cdot \int_{-\infty}^{\infty} (\sigma_\nu x + \mathcal{G}_r(\boldsymbol{\mu}_p) - \mathbb{E}[z_r]) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\
&\quad \cdot \int_{-\infty}^{\infty} (\sigma_\nu y + \mathcal{G}_i(\boldsymbol{\mu}_p) - \mathbb{E}[z_i]) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p ((\mathcal{G}_r(\boldsymbol{\mu}_p) - \mathbb{E}[z_r])(\mathcal{G}_i(\boldsymbol{\mu}_p) - \mathbb{E}[z_i])) \\
&= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r(\boldsymbol{\mu}_p) \mathcal{G}_i(\boldsymbol{\mu}_p) - \mathbb{E}[z_r] \mathbb{E}[z_i]
\end{aligned} \tag{28}$$

Thus, the numerical approximation to the covariance matrix $\boldsymbol{\Sigma}_z$ is then

$$\mathbf{\Sigma}_z = \begin{bmatrix} \sigma_{rr} & \sigma_{ri} \\ \sigma_{ir} & \sigma_{ii} \end{bmatrix} \quad (29)$$

$$\begin{aligned} \sigma_{rr} &= \pi^{-N/2} \sigma_\nu^2 + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_r])^2 \\ \sigma_{ii} &= \pi^{-N/2} \sigma_\nu^2 + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_i^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_i])^2 \\ \sigma_{ri} &= \sigma_{ir} \\ &= \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r(\boldsymbol{\mu}_p) \mathcal{G}_i(\boldsymbol{\mu}_p) - \mathbb{E}[z_r] \mathbb{E}[z_i] \end{aligned}$$

The determinant of the covariance matrix is easily calculated.

$$\begin{aligned} |\mathbf{\Sigma}_z| &= \left(\pi^{N/2} \sigma_\nu^2 \right. \\ &\quad \left. + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_r])^2 \right) \cdot \left(\pi^{N/2} \sigma_\nu^2 + \pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_i^2(\boldsymbol{\mu}_p) - (\mathbb{E}[z_i])^2 \right) \\ &\quad - \left(\pi^{-N/2} \sum_{p=1}^{Q^N} \omega_p \mathcal{G}_r(\boldsymbol{\mu}_p) \mathcal{G}_i(\boldsymbol{\mu}_p) - \mathbb{E}[z_r] \mathbb{E}[z_i] \right)^2 \end{aligned} \quad (30)$$

The mutual information calculation is straightforward once the value of $|\mathbf{\Sigma}_z|$ is known.

$$I(\boldsymbol{\mu}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z} | \boldsymbol{\mu}) = \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\mathbf{\Sigma}_z| \right) - \frac{1}{2} \ln \left((2\pi e)^2 \cdot |\mathbf{\Sigma}_\nu| \right) \quad (31)$$