Notes: Correlation & Mutual Information in Gaussian Channels

M. Rule

May 13, 2011

For a jointly Gaussian pair of random variables, correlation, root mean squared error, correlation, and signal to noise ratio, are all equivalent and can be computed from each-other.

Some identities

Consider two time series x and y that are jointly Gaussian. To simplify things, let x and y have zero mean and unit variance (the math still works out without this assumption, but its also easy to ensure by z-scoring the data). Also, let n be a zero-mean unit-variance Gaussian random variable that captures noise, i.e. fluctuation in y that cannot be explained by x.

Let's say we're interested in a linear relationship between x and y:

$$y = ax + bn.$$

The linear dependence of y on x is summarized by a single parameter

Since the signal and noise are independent, their variances combine linearly:

$$\sigma_y^2 = a^2\sigma_x^2 + b^2\sigma_n^2.$$

The sum $a^2 + b^2$ is constrained by the variances in x, y, and n. In this example we've assumed these are all 1, so

$$a^2 + b^2 = 1.$$

Incorporate this constraint by defining $\alpha = a^2$ and writing

$$\sigma_y^2 = \alpha\sigma_x^2 + (1-\alpha)\sigma_n^2$$

and

$$y = x\sqrt{\alpha} + n\sqrt{1-\alpha}.$$

(We'll show later that α is the squared Pearson correlation coefficient, i.e. it is the coefficient of determination.)

From this the signal-to-noise ratio and mutual information can be calculated

The Signal-to-Noise Ratio (SNR) is the ratio of the signal and noise contributions to x, and simplifies as

$$\mathrm{SNR} = \frac{\sigma_{ax}^2}{\sigma_{bn}^2} = \frac{\alpha\sigma_x^2}{(1-\alpha)\sigma_n^2} = \frac{\alpha}{1-\alpha}.$$

On jointly Gaussian channels mutual information I (in bits, is using $\log_2$) is a monotonic function of SNR, and simplifies as:

$$I = \frac{1}{2}\log_2(1+\mathrm{SNR}) = \frac{1}{2}\log_2\frac{\sigma_y^2}{\sigma_{bn}^2} = \frac{1}{2}\log_2\frac{\sigma_y^2}{(1-\alpha)\sigma_n^2} = \frac{1}{2}\log_2\frac{1}{1-\alpha}.$$

Relationship between a, b, $alpha$, and Pearson correlation ρ Since x and n are independent, the samples of x and n can be viewed as an orthonormal basis for the samples of y, with weights a and b, respectively. This relates the gain parameters to correlation: the tangent of the angle between y and x is just ratio of the noise gain b to the signal gain a:

$$\tan(\theta) = \frac{b}{a} = \frac{\sqrt{1-\alpha}}{\sqrt{\alpha}}$$

Then, $\tan(\theta)$ can be expressed in terms of the correlation coefficient ρ:

$$tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} = \frac{\sqrt{1-\cos(\theta)^2}}{\cos(\theta)} = \frac{\sqrt{1-\rho^2}}{\rho}$$

This implies that

$$\frac{\sqrt{1-\alpha}}{\sqrt{\alpha}} = \frac{\sqrt{1-\rho^2}}{\rho},$$

which implies that that $\alpha = \rho^2$, i.e. $a = \rho$.

A few more identities

This can be used to relate correlation ρ to SNR and mutual information:

$$\mathrm{SNR} = \frac{\rho^2}{1-\rho^2}$$

$$I = \frac{1}{2}\log_2 \frac{1}{1-\rho^2} = -\frac{1}{2}\log_2(1-\rho^2)$$

If $\phi = \sqrt{1-\rho^2}$ is the correlation of y and the noise n (i.e. ϕ is the amplitude of the noise contribution to y), then information is simply $I = -\log_2(\phi)$.

Mean squared error (MSE) is also related :

$$\text{MSE} = (1-\rho)^2 + (1-\rho^2) = 1 - 2\rho + 1 = 2(1-\rho),$$

which implies that

$$\rho = 1 - \frac{1}{2}\text{MSE},$$

and gives a relationship between mutual information and mean squared error:

$$I = -\frac{1}{2}lg(1-\rho^2) = -\frac{1}{2}\log_2(1-(1-\text{MSE}/2)^2)$$

These relationships between correlation ρ, mean squared error, mutual information, and signal to noise ratio, all increase monotonically. They all summarize the relatedness of x and y. For purposes, e.g. of ranking a collection of x in terms of how much they tell us about y, they are equivalent.