Analysis of Uncertainties **Fall 2014**

Lecture 4: 31 October 2014

Lecturer: Prof. Kate Scholberg *Notes by: Douglas Davis*

4.1 Error matrices continued

Last lecture we introduced error matrices. Let's briefly remind ourselves of error propagation for some measurement x which is a function of many variables $u, v, w, \ldots$.

$$\sigma_x^2 = \sigma_u^2 \left(\frac{\partial x}{\partial u}\right) + \sigma_v^2 \left(\frac{\partial x}{\partial v}\right) + \ldots + 2\sigma_{uv}^2 \left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right) + \ldots \tag{4.1}$$

where

$$\sigma_{uv}^2 = \lim_{N\to\infty} \frac{1}{N} \sum_i^N \left(u_i - \overline{u}\right)\left(v_i - \overline{v}\right) = \left\langle \left(u - \overline{u}\right)\left(v - \overline{v}\right)\right\rangle. \tag{4.2}$$

If u and v are uncorrelated, σ_{uv} is of course zero. If u and v are correlated – we have two possible situations, *positive correlation* (or simply *correlated* variables) and *negative correlation* (or simply *anticorrelated* variables). For two variables u and v, positive correlation would correspond to u and v growing or decreasing together – an example would be height and weight (with a large sample, the majority of taller people would weight more than those shorter than them). Negative correlation between u and v is a situation where as one grows the other decreases. Take the height of a horse racing jockey and the speed the horse will run. Figure 4.1 shows samples with their respective correlations. The covariance σ_{uv}^2 has dimensions – it's often

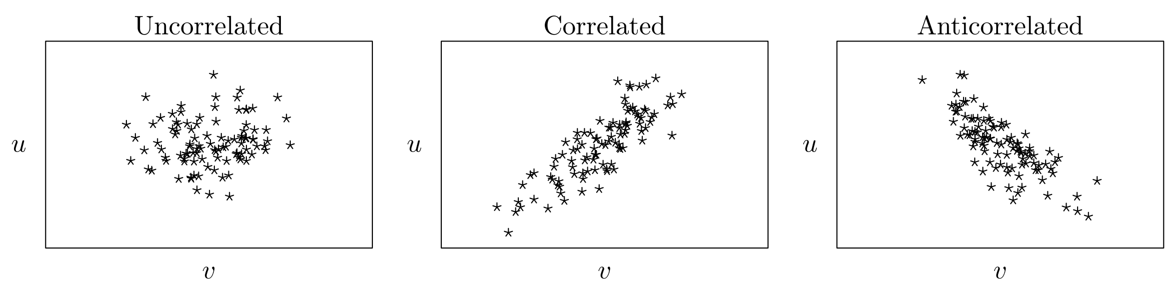

Figure 4.1: Three samples which correspond to uncorrelated variables, correlated variables, and uncorrelated variables.

desired to have a dimensionless quantity. For this we will define the correlation coefficient:

$$\rho = \frac{\sigma_{uv}^2}{\sigma_u \sigma_v} = \frac{\mathrm{cov}(u, v)}{\sigma_u \sigma_v} \tag{4.3}$$

It's clear that ρ is bounded: $-1 \le \rho \le 1$. We can define three levels of correlation in table 4.1.

Now let's dive back into error ellipses. Take two uncorrelated outcomes u and v with Gaussian probabilities:

$$P(u) = \frac{1}{\sqrt{2\pi}\sigma_u} e^{-u^2/2\sigma_u^2} \quad P(v) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-v^2/2\sigma_v^2}. \tag{4.4}$$

Table 4.1: ρ values table

$\rho = 0$	uncorrelated
$\rho = 1$	complete correlation
$\rho = -1$	complete anticorrelation

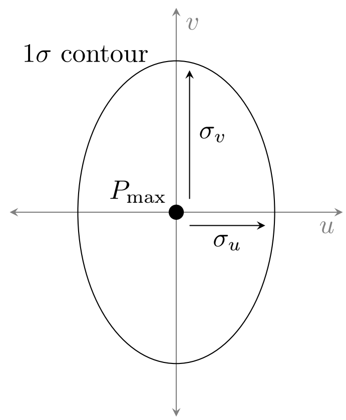

Figure 4.2: The 1σ contour displayed in an error ellipse corresponding to a pair of uncorrelated variables

It follows that

$$P(u)P(v) = \frac{1}{2\pi\sigma_u\sigma_v} \exp\left[-\frac{1}{2}\left(\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right]. \tag{4.5}$$

We can write that $P = P_{\max}e^{-1/2}$ when

$$\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} = 1.$$

This is the equation of an ellipse where the major and minor axes are given by the the square root of the variance of each quantity. Figure 4.2 shows this in graphical form. The volume inside the 1σ contour holds approximately 68% of the totla probability. We can write equation 4.5 in a new form using vectors and matrices:

$$P(u,v) = \frac{1}{2\pi\left|\mathbf{M}\right|^{1/2}} \exp\left[-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{x}\right]. \tag{4.6}$$

The term inside the exponential ($\mathbf{x}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{x}$) can be written:

$$\begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} 1/\sigma_u^2 & 0 \\ 0 & 1/\sigma_v^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

From here we define the error matrix $\mathbf{M}$ (for uncorrelated variables):

$$\mathbf{M} = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}. \tag{4.7}$$

In generaal (or say for possibly correlated variables) the matrix takes the general form:

$$M_{ij} = \left\langle \left(u_i - \overline{u}_i\right)\left(u_j - \overline{u}_j\right)\right\rangle, \tag{4.8}$$

where u_i corresponds to some variable set $\{u, v, w, \dots\}$. Using the numbered indices again, we can write an arbitrary error matrix in the form:

$$\mathbf{M} = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}(1,2) & \mathrm{cov}(1,3) & \dots \\ \mathrm{cov}(1,2) & \sigma_2^2 & \mathrm{cov}(2,3) & \dots \\ \mathrm{cov}(1,3) & \mathrm{cov}(2,3) & \sigma_3^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.9}$$

Therefore if we go back to our standard two variable correlated form (using u and v again) we would write the error matrix in the form:

$$\mathbf{M} = \begin{pmatrix} \sigma_u^2 & \mathrm{cov}(u,v) \\ \mathrm{cov}(u,v) & \sigma_v^2 \end{pmatrix}. \tag{4.10}$$

Using the definition of the correlation coefficient, we can now write the probability equation (with the ellipse equation in the exponential) in the following form:

$$P(u,v) = \frac{1}{2\pi\sigma_u\sigma_v}\frac{1}{\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2}\left[\frac{1}{(1-\rho)^2}\left(\frac{u^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}-\frac{2\rho uv}{\sigma_u\sigma_v}\right)\right]\right\}. \tag{4.11}$$

This is assuming a mean of zero for u and v, for non zero means we simply change all u and v terms to $(u-\overline{u})$ and $(v-\overline{v})$, respectively.

Now we'll generalize to an arbitrary number of variables (here we say k variables):

$$P(x_1, x_2, \dots, x_k) = \frac{1}{(2\pi)^{k/2}}\frac{1}{|\mathbf{M}|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\mathbf{M}(\mathbf{x}-\boldsymbol{\mu})\right], \tag{4.12}$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} \quad \text{or} \quad \boldsymbol{\mu} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \overline{x}_k \end{pmatrix}. \tag{4.13}$$

We note here that the eigenvalues of the error matrix are the semi axes of the error ellipse.

Make diagram showing error ellipse

Now that we've done a pretty in depth discussion of the error matrix let's talk about using it. Let's say we have a variable which is a function of measurable quantities with known uncertainties – we can use the error matrix for those values to determine the uncertainties in our desired variable. Our desired uncertainty is the uncertainty in some parameter ζ, and it is a function of N variables $\{x_1, \dots, x_N\}$:

$$\zeta = f(x_1, x_2, \dots, x_N).$$

We can define vector $\mathbf{D}$ of derivates of f w.r.t. all independent variables:

$$\mathbf{D} = \begin{pmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_N \end{pmatrix} \tag{4.14}$$

Then the variance in ζ is defined as:

$$\sigma_\zeta^2 = \mathbf{D}^{\mathrm{T}}\mathbf{M}\mathbf{D} \tag{4.15}$$

For variable transformations – for example if x is a function of r and θ and y is a function of r and θ, then we develop a Jacobi style matrix to determine an error matrix for x and y systems:

$$\mathbf{A} = \begin{pmatrix} \partial x/\partial r & \partial y/\partial r \\ \partial x/\partial \theta & \partial y/\partial \theta \end{pmatrix}. \tag{4.16}$$

We then define the new error matrix (for x and y) as:

$$\mathbf{M}_{xy} = \mathbf{A}^{\mathrm{T}}\mathbf{M}_{r\theta}\mathbf{A} \tag{4.17}$$

This of course can be generalized to many variables. For example if $\{\alpha_1, \ldots, \alpha_N\}$ are dependent on $\{\beta_1, \ldots, \beta_M\}$, then $\mathbf{A}$ would have the form:

$$\mathbf{A} = \begin{pmatrix} \partial\alpha_1/\partial\beta_1 & \partial\alpha_2/\partial\beta_1 & \ldots & \partial\alpha_N/\partial\beta_1 \\ \partial\alpha_1/\partial\beta_2 & \partial\alpha_2/\partial\beta_2 & \ldots & \partial\alpha_N/\partial\beta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial\alpha_1/\partial\beta_M & \partial\alpha_2/\partial\beta_M & \ldots & \partial\alpha_N/\partial\beta_M \end{pmatrix} \tag{4.18}$$

Now if we go back to the (x, y) and (r, θ) example, we can define some variable z which is a function of x and y:

$$z = f(x, y)$$

And now to determine the uncertainty on a variable in some transformed coordinates we develop another derivative vector in the new system (call it the prime system):

$$\mathbf{D}' = \begin{pmatrix} \partial z/\partial x \\ \partial z/\partial y \end{pmatrix} \tag{4.19}$$

And we calculate the variance of z as:

$$\sigma_z^2 = \mathbf{D}'^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{M}_{r\theta}\mathbf{A}\mathbf{D}', \tag{4.20}$$

which would just be (as intuitively expected based on equation 4.15):

$$\sigma_z^2 = \mathbf{D}'^{\mathrm{T}}\mathbf{M}_{xy}\mathbf{D}' \tag{4.21}$$

And of course this can be generalized back to our α and β system if we define some ξ as a function of all of the α's. Then:

$$\mathbf{D}' = \begin{pmatrix} \partial\xi/\partial\alpha_1 \\ \partial\xi/\partial\alpha_2 \\ \vdots \\ \partial\xi/\partial\alpha_N \end{pmatrix} \tag{4.22}$$