CONFIDENCE REGION RADIUS

Multivariate Normal Distribution

The multivariate normal distribution (Altham 2006; NIST 2006) has probability density function

$$f(\boldsymbol{x}) = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]}{\sqrt{(2\pi)^n |\boldsymbol{V}|}} \tag{1}$$

where n is the number of dimensions, $\boldsymbol{x}$ is a column vector of random variables, $\boldsymbol{\mu}$ is a column vector of corresponding means, and $\boldsymbol{V}$ is the covariance matrix.

Because we wish to integrate $f(\boldsymbol{x})$ over an n-dimensional ball centered at $\boldsymbol{\mu}$, we may, without loss of generality, choose a coordinate system with origin $\boldsymbol{\mu}$ and axis directions given by the eigenvectors of $\boldsymbol{V}$. A three-dimensional example is illustrated in Fig. 1. In this new coordinate system, the random variables are uncorrelated with zero mean, and the covariance matrix is diagonal with variances given by the eigenvalues of $\boldsymbol{V}$. Moreover, because $f(\boldsymbol{x})$ is symmetric, we may choose any computationally convenient permutation of the variances.

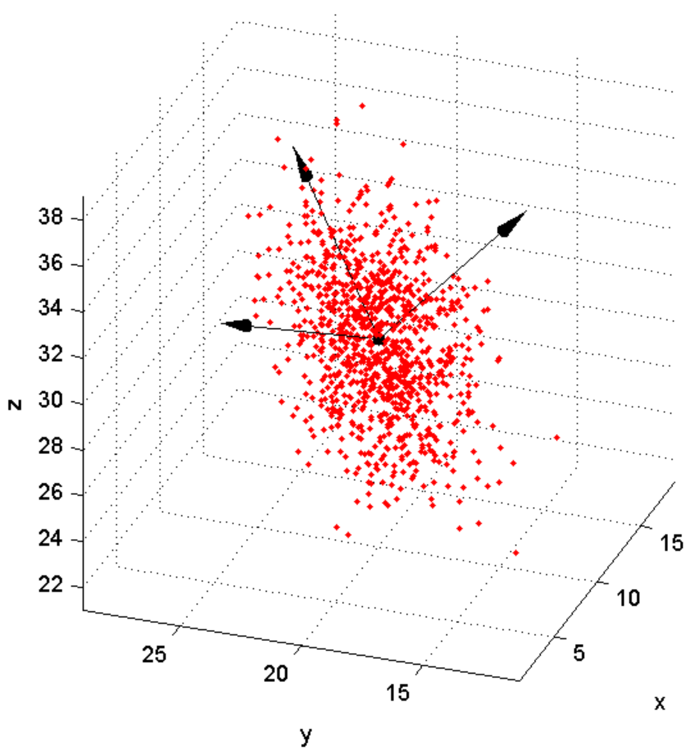

Figure 1. New Coordinate System

Univariate Normal Distribution

For a normally distributed variable with zero mean, the probability that an observation will be contained in the interval [-R,R] is

$$p = \frac{2}{\sigma\sqrt{2\pi}} \int_0^R \exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right) dx \,. \tag{2}$$

Evaluating (2)

$$p = \operatorname{erf}\left(\frac{1}{\sqrt{2}}\frac{R}{\sigma}\right) \Rightarrow R = \sqrt{2}\,\sigma\,\operatorname{erf}^{-1}(p) \tag{3}$$

where erf is the error function.

Bivariate Normal Distribution

For uncorrelated, normally distributed variables with zero mean, the probability that an observation will be contained in area A is

$$p = \frac{1}{\sigma_x\sigma_y 2\pi} \iint_A \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] dA \,. \tag{4}$$

Casting (4) in polar coordinates and integrating over the disk centered at the origin with radius R

$$p = \frac{4}{\sigma_x\sigma_y 2\pi} \int_0^R \int_0^{\frac{\pi}{2}} r \exp\left[-\frac{r^2}{2}\left(\frac{\cos^2\theta}{\sigma_x^2} + \frac{\sin^2\theta}{\sigma_y^2}\right)\right] d\theta\, dr \,. \tag{5}$$

Substituting $\sin^2\theta = 1 - \cos^2\theta$

$$p = \frac{2}{\sigma_x\sigma_y\pi} \int_0^R r \exp\left(-\frac{1}{2}\frac{r^2}{\sigma_y^2}\right) \int_0^{\frac{\pi}{2}} \exp(-2\,a\,r^2\cos^2\theta)\, d\theta\, dr \tag{6}$$

where $a = (1/\sigma_x^2 - 1/\sigma_y^2)/4$.

Evaluating the inner integral of (6)

$$p = \frac{1}{\sigma_x\sigma_y} \int_0^R r \exp(b\,r^2)\; I_0(a\,r^2)\, dr \tag{7}$$

where $b = -(1/\sigma_x^2 + 1/\sigma_y^2)/4$, and I_0 is the zero order modified Bessel function of the first kind.

Substituting $I_0(a\,r^2) = \sum_{k=0}^{\infty} \frac{1}{(k!)^2} \left(\frac{a\,r^2}{2} \right)^{2k}$ (Abramowitz and Stegun 1972)

$$p = \frac{1}{\sigma_x \sigma_y} \sum_{k=0}^{\infty} \frac{1}{(k!)^2} \left(\frac{a}{2} \right)^{2k} \int_0^R r^{4k+1} \exp(b\,r^2)\,dr\,. \tag{8}$$

Integrating (8) term-by-term

$$p = \frac{1}{a\,\sigma_x \sigma_y} \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2} \left(-\frac{1}{2}\frac{a}{b} \right)^{2k+1} P(2k+1, -bR^2) \tag{9}$$

where the central binomial coefficients (Sloane 2006) $f(k) = (2k)!/(k!)^2$ are computed via the recurrence relation $f(k) = f(k-1)(4 - 2/k)$, and P is the incomplete gamma function (Abramowitz and Stegun 1972).

When $\sigma_x = \sigma_y = \sigma$

$$p = 1 - \exp\left(-\frac{1}{2}\frac{R^2}{\sigma^2} \right) \Rightarrow R = \sigma\sqrt{-2\ln(1-p)}\,. \tag{10}$$

Trivariate Normal Distribution

For uncorrelated, normally distributed variables with zero mean, the probability that an observation will be contained in volume V is

$$p = \frac{1}{\sigma_x \sigma_y \sigma_z (2\pi)^{3/2}} \iiint_V \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2} \right) \right] dV\,. \tag{11}$$

Casting (11) in cylindrical coordinates and integrating over the ball centered at the origin with radius R

$$p = \frac{8\int_0^R \int_0^{\frac{\pi}{2}} r \exp\left[-\frac{r^2}{2}\left(\frac{\cos^2\theta}{\sigma_x^2} + \frac{\sin^2\theta}{\sigma_y^2} \right) \right] \int_0^{\sqrt{R^2-r^2}} \exp\left(-\frac{1}{2}\frac{z^2}{\sigma_z^2} \right) dz\,d\theta\,dr}{\sigma_x \sigma_y \sigma_z (2\pi)^{3/2}}\,. \tag{12}$$

Evaluating the innermost integral of (12)

$$p = \frac{2}{\sigma_x \sigma_y \pi} \int_0^R \int_0^{\frac{\pi}{2}} r \exp\left[-\frac{r^2}{2} \left(\frac{\cos^2 \theta}{\sigma_x^2} + \frac{\sin^2 \theta}{\sigma_y^2} \right) \right] \operatorname{erf}\left(\frac{\sqrt{R^2 - r^2}}{\sqrt{2}\, \sigma_z} \right) d\theta \, dr \,. \tag{13}$$

Substituting $\sin^2 \theta = 1 - \cos^2 \theta$

$$p = \frac{2}{\sigma_x \sigma_y \pi} \int_0^R r \operatorname{erf}\left(\frac{\sqrt{R^2 - r^2}}{\sqrt{2}\, \sigma_z} \right) \exp\left(-\frac{1}{2} \frac{r^2}{\sigma_y^2} \right) \int_0^{\frac{\pi}{2}} \exp(-2\, a\, r^2 \cos^2 \theta) \, d\theta \, dr \tag{14}$$

where $a = (1/\sigma_x^2 - 1/\sigma_y^2)/4$.

Evaluating the inner integral of (14)

$$p = \frac{1}{\sigma_x \sigma_y} \int_0^R r \operatorname{erf}\left(\frac{\sqrt{R^2 - r^2}}{\sqrt{2}\, \sigma_z} \right) \exp(b\, r^2) \;\; I_0(a\, r^2) \, dr \tag{15}$$

where $b = -(1/\sigma_x^2 + 1/\sigma_y^2)/4$.

When $\sigma_x = \sigma_y = \sigma_h$

$$p = \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{R}{\sigma_z} \right) - \frac{1}{c} \exp\left(-\frac{1}{2} \frac{R^2}{\sigma_h^2} \right) \operatorname{erf}\left(\frac{c}{\sqrt{2}} \frac{R}{\sigma_z} \right) \tag{16}$$

where $c = \sqrt{1 - \sigma_z^2 / \sigma_h^2}$ is either real or pure imaginary.

When $\sigma_x = \sigma_y = \sigma_z = \sigma$

$$p = \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{R}{\sigma} \right) - \sqrt{\frac{2}{\pi}} \frac{R}{\sigma} \exp\left(-\frac{1}{2} \frac{R^2}{\sigma^2} \right). \tag{17}$$

Applying Newton's method to (17), R may be determined using the iteration

$$R_{k+1} = R_k + \frac{\sigma^2}{R_k} + \sqrt{\frac{\pi}{2}} \frac{\sigma^3}{R_k^2} \exp\left(\frac{1}{2} \frac{R_k^2}{\sigma^2} \right) \left[p - \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{R_k}{\sigma} \right) \right] \tag{18}$$

with an initial estimate given by (10).

Davis and Kleder (2006) solve (7), (9), (15), (16), and (17) for radius R corresponding to probability p. The algorithm is an extension of code presented by Kleder (2004) with enhancements suggested by Koopman (2004).

References

Abramowitz, M. and Stegun, I. A., eds. (1972). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards. 6.5.1, 7.1.23, 7.1.29, 9.6.12. (http://www.math.sfu.ca/~cbm/aands/frameindex.htm)

Altham, P. M. E. (2006). "Applied multivariate analysis, notes." (http://www.statslab.cam.ac.uk/~pat/AppMultNotes.pdf)

Davis, T. and Kleder, M. (2006). “Confidence region radius.” Mathworks Central File Exchange. (http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=10526&objectType=file)

Kleder, M. (2004). "An algorithm for converting covariance to spherical error probable." Mathworks Central File Exchange. (http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=5688&objectType=file)

Koopman, R. (2004). "Joint normal distribution integral over the unit sphere." The Math Forum. (http://mathforum.org/kb/message.jspa?messageID=1552353&tstart=0)

National Institute of Standards and Technology (NIST 2006). "The multivariate normal distribution." (http://www.itl.nist.gov/div898/handbook/pmc/section5/pmc542.htm)

Sloane, N. J. A. (2006). “The on-line encyclopedia of integer sequences.” AT&T Labs Research. A000984 (http://www.research.att.com/~njas/sequences/A000984)