Ellipses made Useful

Author: Han de Bruijn
Dated: May 2005
Updated: June 2006

Skewed 2-D Bell Shape

Associated with the first and second order moments in one dimension is the Gauss function, also known as the *normal distribution* in Statistics:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} = \frac{1}{\sqrt{2\pi\sigma_{xx}}} e^{-\frac{1}{2}(x-\mu_x)^2/\sigma_{xx}}$$

The exponent (apart from the factor 1/2) could have been written as:

$$(x-\mu_x)^2/\sigma_{xx} = (x-\mu_x)\frac{1}{\sigma_{xx}}(x-\mu_x)$$

In the general two-dimensional case, σ_{xx} will be replaced by the tensor:

$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix}$$

And the inverse $1/\sigma_{xx}$ by the inverse of this matrix:

$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix}^{-1} = \begin{bmatrix} \sigma_{yy} & -\sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{bmatrix} /(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)$$

The accompanying quadratic form is:

$$\begin{bmatrix} (x-\mu_x) & (y-\mu_y) \end{bmatrix} \begin{bmatrix} \sigma_{yy} & -\sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{bmatrix} \begin{bmatrix} (x-\mu_x) \\ (y-\mu_y) \end{bmatrix} /(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)$$

$$= \frac{\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}$$

This in turn corresponds to the generalization of the Gauss Function in 2-D:

$$g(x,y) = e^{-\frac{1}{2}\left[\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2\right]/(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2)}$$

A simplified quadratic form for the inverse problem can be found easily, because the eigenvalues of an inverse matrix are always the inverses of the eigenvalues of the original tensor. The latter are λ_1 and λ_2. Hence the former are found immediately to be:

$$1/\lambda_1 \quad \text{and} \quad 1/\lambda_2$$

This in turn means that the Gauss function, when transformed to eigenvector coordinates, is simply given by:

$$g(x,y) = e^{-\frac{1}{2}\left[(x-\mu_x)^2/\lambda_1 + (y-\mu_y)^2/\lambda_2\right]}$$

What's still missing is a *norming factor* for the skewed 2-D Gaussian function. To this end, integrate the function $g(x,y)$ over the whole plane:

$$\iint g(x,y)\,dx\,dy = \iint e^{-\frac{1}{2}\left[(x-\mu_x)^2/\lambda_1 + (y-\mu_y)^2/\lambda_2\right]}\,dx\,dy$$

Substitute $u = (x-\mu_x)/\sqrt{\lambda_1}$ and $v = (y-\mu_y)/\sqrt{\lambda_2}$:

$$= \iint e^{-\frac{1}{2}(u^2+v^2)}\,d(u\sqrt{\lambda_1})\,d(v\sqrt{\lambda_2}) = \sqrt{\lambda_1\lambda_2}\iint e^{-\frac{1}{2}(u^2+v^2)}\,du\,dv =$$

Transform to polar coordinates:

$$= \sqrt{\lambda_1\lambda_2}\iint e^{-\frac{1}{2}r^2}\,rdrd\phi = -\sqrt{\lambda_1\lambda_2}\iint e^{-\frac{1}{2}r^2}\,d(-\frac{1}{2}r^2)d\phi =$$

$$= -\sqrt{\lambda_1\lambda_2}\,2\pi\left[e^{-\frac{1}{2}r^2}\right]_0^{\infty} = \sqrt{\lambda_1\lambda_2}\,2\pi = 2\pi\,\sqrt{Det}$$

Thus the norming factor for the 2-D skewed Gaussian function is, when used over the whole plane from $-\infty$ to $+\infty$:

$$2\pi\sqrt{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}$$

The normed (and skewed) two-dimensional Gaussian distribution function is completed herewith as:

$$g(x,y) = \frac{e^{-\frac{1}{2}\left[\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2\right]/(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2)}}{2\pi\sqrt{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}}$$

A rather obvious generalization of this, to n-dimensional space, is mentioned as equation (1.8) in the Introduction of the book *Neuro-Fuzzy Pattern Recognition* by Sankar K. Pal and Susmita Mitra (1999, John Wiley & Sons, Inc.):

$$p(\vec{x}|C_j) = \frac{1}{(2\pi)^{n/2}\left|\Sigma_j\right|^{1/2}} exp\left[-\frac{1}{2}(\vec{x}-\vec{m}_j)'\Sigma_j^{-1}(\vec{x}-\vec{m}_j)\right]$$

Drawing an Ellipse

The following general expression will be proposed for the ellipse of inertia:

$$\frac{\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2} = 1$$

When making a drawing of such an ellipse, it would be more handsome to have it in parametrized form. That is, we seek an equivalent like this:

$$\begin{aligned} x &= \mu_x + a_x\, cos(t) + a_y\, sin(t) \\ y &= \mu_y + b_x\, cos(t) + b_y\, sin(t) \end{aligned}$$

Multiply the first equation with b_x, the second with a_x and substract:

$$b_x(x-\mu_x) - a_x(y-\mu_y) = (a_y b_x - a_x b_y)\, sin(t) \quad \Longrightarrow$$

$$sin(t) = \frac{b_x(x-\mu_x) - a_x(y-\mu_y)}{a_y b_x - a_x b_y}$$

Multiply the first equation with b_y, the second with a_y and substract:

$$b_y(x-\mu_x) - a_y(y-\mu_y) = (b_y a_x - b_x a_y)\, cos(t) \quad \Longrightarrow$$

$$cos(t) = \frac{b_y(x-\mu_x) - a_y(y-\mu_y)}{b_y a_x - b_x a_y}$$

Now use the well known identity:

$$cos^2(t) + sin^2(t) = 1$$

Giving:

$$\left(\frac{b_x x' - a_x y'}{b_y a_x - b_x a_y}\right)^2 + \left(\frac{b_y x' - a_y y'}{a_y b_x - a_x b_y}\right)^2 = 1$$

Where $x' = x - \mu_x$ and $y' = y - \mu_y$. This result must be comparable with:

$$\frac{\sigma_{yy}(x')^2 - 2\sigma_{xy}x'y' + \sigma_{xx}(y')^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2} = 1$$

Drop the primes $'$ in both equations for the sake of simplicity. Look after:

$$\left(\frac{b_x x - a_x y}{b_y a_x - b_x a_y}\right)^2 + \left(\frac{b_y x - a_y y}{a_y b_x - a_x b_y}\right)^2 = 1$$

And work out:

$$\frac{(b_x^2 + b_y^2)\, x^2 - 2(a_x b_x + a_y b_y)\, xy + (a_x^2 + a_y^2)\, y^2}{(b_y a_x - b_x a_y)^2} = 1$$

As compared with:

$$\frac{\sigma_{yy}x^2 - 2\sigma_{xy}xy + \sigma_{xx}y^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2} = 1$$

We thus find:

$$a_x^2 + a_y^2 = \sigma_{xx} \quad \text{and} \quad b_x^2 + b_y^2 = \sigma_{yy} \quad \text{and} \quad a_x b_x + a_y b_y = \sigma_{xy}$$

It also follows that the following must hold:

$$\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 = (a_x b_y - b_x a_y)^2$$

Check this:

$$(a_x^2 + a_y^2)(b_x^2 + b_y^2) - (a_x b_x + a_y b_y)^2 =$$
$$a_x^2 b_x^2 + a_x^2 b_y^2 + a_y^2 b_x^2 + a_y^2 b_y^2 - (a_x^2 b_x^2 + a_y^2 b_y^2 + 2a_x b_x a_y b_y) =$$
$$(a_x b_y)^2 + (b_x a_y)^2 - 2(a_x b_y)(b_x a_y) = (a_x b_y - b_x a_y)^2$$

Happy maybe that this is consistent, we are nevertheless faced with a little problem: there are four unknowns (a_x, a_y, b_x, b_y), but only three equations to solve for them. This problem is readily tackled by the meaning of the equations, though. The first equation means that the length of the vector $\vec{a}$ is equal to $\sqrt{\sigma_{xx}}$. The second equation means that the length of the vector $\vec{b}$ is equal to $\sqrt{\sigma_{yy}}$. The third equation means that the cosine of the angle between $\vec{a}$ and $\vec{b}$ is given by the inner product divided by these lengths: $\cos(\theta) = \sigma_{xy}/(\sqrt{\sigma_{xx}\sigma_{yy}})$. Now it is clear that these three facts are quite independent of an eventual rotation of the coordinate system:

$$\begin{aligned} a'_x &= +\cos(\tau)\ a_x + \sin(\tau)\ a_y \\ a'_y &= -\sin(\tau)\ a_x + \cos(\tau)\ a_y \end{aligned}$$

When applied to the parametrized equations, for example the first of the two:

$$x - \mu_x = a'_x\ cos(t) + a'_y\ sin(t) =$$
$$[cos(\tau)\ a_x + sin(\tau)\ a_y]\, cos(t) + [-sin(\tau)\ a_x + cos(\tau)\ a_y]\, sin(t) =$$
$$[cos(\tau)cos(t) - sin(\tau)sin(t)]\, a_x + [sin(\tau)cos(t) + cos(\tau)sin(t)]\, a_y =$$
$$cos(t + \tau)\ a_x + sin(t + \tau)\ a_y$$

We conclude that a change of coordinate system is just the same as a different offset for the running parameter. Nothing interesting actually. This the reason why we are free to replace the four unknowns by only three. For example, choose $\vec{a}$ in the same direction as the x-axis, meaning that $a_y = 0$. Then the solution of our problem becomes:

$$a_x^2 = \sigma_{xx} \quad \Longrightarrow \quad a_x = \sqrt{\sigma_{xx}} \quad \text{and} \quad a_x b_x = \sigma_{xy} \quad \Longrightarrow \quad b_x = \sigma_{xy}/\sqrt{\sigma_{xx}}$$

$$b_x^2 + b_y^2 = \sigma_{yy} \quad \Longrightarrow \quad b_y = \sqrt{\sigma_{yy} - \sigma_{xy}^2/\sigma_{xx}} = \sqrt{\frac{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}{\sigma_{xx}}}$$

Intersecting an Ellipse

The following expression will be proposed for the bounding ellipse of a curve:

$$\sigma_{yy}(x-\mu_x)^2 - 2\,\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2 - E = 0$$

Bounding ellipses are employed in our programs for clipping complete contours "at first sight", that is: without having to traverse all vertices of these contours. For this purpose, they are far more handsome than bounding boxes.
In order to make this clipping machinery operational, we have to answer the following question. What are the intersection points of a bounding ellipse with an arbitrary straight line segment:

$$\begin{array}{c} x = x_1 + \lambda(x_2 - x_1) \\ y = y_1 + \lambda(y_2 - y_1) \end{array} \quad \text{where} \quad 0 \le \lambda \le 1$$

Substitute x and y in the equation for the ellipse:

$$\begin{gathered} \sigma_{yy}\left[\lambda(x_2-x_1)+(x_1-\mu_x)\right]^2 \\ -2\,\sigma_{xy}\left[\lambda(x_2-x_1)+(x_1-\mu_x)\right]\left[\lambda(y_2-y_1)+(y_1-\mu_y)\right] \\ +\sigma_{xx}\left[\lambda(y_2-y_1)+(y_1-\mu_y)\right]^2 - E = 0 \end{gathered}$$

And work out:

$$\begin{gathered} \left[\sigma_{yy}(x_2-x_1)^2 - 2\,\sigma_{xy}(x_2-x_1)(y_2-y_1) + \sigma_{xx}(y_2-y_1)^2\right]\lambda^2 \\ +\,2\left[\sigma_{yy}(x_2-x_1)(x_1-\mu_x) - \sigma_{xy}(x_2-x_1)(y_1-\mu_y)\right. \\ \left.-\sigma_{xy}(y_2-y_1)(x_1-\mu_x) + \sigma_{xx}(y_2-y_1)(y_1-\mu_y)\right]\lambda \\ +\left[\sigma_{yy}(x_1-\mu_x)^2 - 2\,\sigma_{xy}(x_1-\mu_x)(y_1-\mu_y) + \sigma_{xx}(y_1-\mu_y)^2\right] - E = 0 \end{gathered}$$

This is a quadratic equation:

$$A\lambda^2 + 2\,B\lambda + C = 0$$

Where:

$$\begin{aligned} A &= \sigma_{yy}x_{21}^2 - 2\,\sigma_{xy}x_{21}y_{21} + \sigma_{xx}y_{21}^2 \\ B &= \sigma_{yy}x_{21}x_{1\mu} - \sigma_{xy}\left[x_{21}y_{1\mu} + y_{21}x_{1\mu}\right] + \sigma_{xx}y_{21}y_{1\mu} \\ C &= \sigma_{yy}x_{1\mu}^2 - 2\,\sigma_{xy}x_{1\mu}y_{1\mu} + \sigma_{xx}y_{1\mu}^2 - E \end{aligned}$$

The following substitutions have been done:

$$x_{21} = x_2 - x_1 \qquad y_{21} = y_2 - y_1 \qquad x_{1\mu} = x_1 - \mu_x \qquad y_{1\mu} = y_1 - \mu_y$$

Whether the ellipse is intersected by the line is determined by the discriminant D in the first place:

$$\begin{array}{lcl} D = B^2 - AC < 0 & \Longrightarrow & \text{no intersections} \\ D = B^2 - AC = 0 & \Longrightarrow & \text{maybe tangent} \\ D = B^2 - AC > 0 & \Longrightarrow & \text{maybe intersecting} \end{array}$$

In case $B^2 - AC = 0$, the tangent point (x, y) is given by:

$$\begin{array}{c} x = x_1 + \lambda(x_2 - x_1) \\ y = y_1 + \lambda(y_2 - y_1) \end{array} \quad \text{where} \quad \lambda = -\frac{B}{A} \quad \text{but only if} \quad 0 \le \lambda \le 1$$

In case $B^2 - AC > 0$, the intersection points (x, y) are given by:

$$\begin{array}{c} x = x_1 + \lambda(x_2 - x_1) \\ y = y_1 + \lambda(y_2 - y_1) \end{array} \quad \text{where} \quad \lambda = \frac{-B \pm \sqrt{D}}{A} \quad \text{but only if} \quad 0 \le \lambda \le 1$$

The above provides (more than) enough material to efficiently implement the clipping of complete contours. For example, if the discriminant of a clipping line and an ellipse is negative, then all vertices of a contour are clipped if only the midpoint (μ_x, μ_y) of its bounding ellipse is clipped.

Bounding Ellipse

When it comes to the preprocessing of domains and contours, Bounding Ellipses are virtually indespensible. They are far more handsome than bounding boxes of whatever kind in the first place. The only quantities to be determined are the midpoint and the second order variances of the body (or its boundary curve) under consideration. A bounding ellipse is then based upon the *inverse* tensor of inertia / matrix of variances. The equation of an ellipse of inertia is given by:

$$\left[\begin{array}{cc} (x - \mu_x) & (y - \mu_y) \end{array} \right] \left[\begin{array}{cc} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{array} \right]^{-1} \left[\begin{array}{c} (x - \mu_x) \\ (y - \mu_y) \end{array} \right] = 1 \quad \Longrightarrow$$

$$\frac{\sigma_{yy}(x - \mu_x)^2 - 2\sigma_{xy}(x - \mu_x)(y - \mu_y) + \sigma_{xx}(y - \mu_y)^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2} = E$$

Where σ are the variances and μ are the midpoint coordinates. A bounding ellipse is defined as an ellipse of inertia which its right hand side modified: the constant 1 is replaced by a another constant E. Here E should be adapted in such a way that the whole boundary of the area of interest is contained inside the ellipse. Thus E is the maximum of:

$$\frac{\sigma_{yy}(x_k - \mu_x)^2 - 2\sigma_{xy}(x_k - \mu_x)(y_k - \mu_y) + \sigma_{xx}(y_k - \mu_y)^2}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2}$$

where (k) runs through all vertices of the boundary contours.

Best Fit Ellipse

Quite another problem is whether and when an ellipse of inertia can be considered as a Best Fit Ellipse. The latter shall be defined with help of a Least Squares minimization principle, as follows:

$$\sum_k w_k \left[\frac{\sigma_{yy}(x_k-\mu_x)^2 - 2\sigma_{xy}(x_k-\mu_x)(y_k-\mu_y) + \sigma_{xx}(y_k-\mu_y)^2}{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2} - E \right]^2$$

$= minimum(E)$. Differentiating to E simply gives:

$$\sum_k w_k \left[\frac{\sigma_{yy}(x_k-\mu_x)^2 - 2\sigma_{xy}(x_k-\mu_x)(y_k-\mu_y) + \sigma_{xx}(y_k-\mu_y)^2}{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2} - E \right] = 0$$

If and only if:

$$\sigma_{yy}\sum_k w_k(x_k-\mu_x)^2 - 2\sigma_{xy}\sum_k w_k(x_k-\mu_x)(y_k-\mu_y) + \sigma_{xx}\sum_k w_k(y_k-\mu_y)^2$$
$$= E(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2)$$

If and only if:

$$\sigma_{yy}\sigma_{xx} - 2\sigma_{xy}\sigma_{xy} + \sigma_{xx}\sigma_{yy} = 2\left(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2\right) = E(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2)$$

If and only if $E = 2$. Thus the equation of the Best Fit Ellipse is:

$$\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2$$
$$= 2\left(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2\right)$$

Or alternatively:

$$\frac{1}{2}\,\frac{\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2}{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2} = 1$$

Extreme Cases

When looking for a sensible meaning of the determinant $\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2$, it can be proved in the first place that:

$$\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2 \ge 0$$

Which is a direct consequence of Schwartz inequality:

$$(\vec{x}\cdot\vec{y})^2 \le (\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y}) \quad \Longleftrightarrow \quad \left(\sum_k w_k x_k y_k\right)^2 \le \left(\sum_k w_k x_k^2\right)\left(\sum_k w_k y_k^2\right)$$

Remember the Ellipse of Inertia:

$$\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2 = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$$

For $\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 = 0$ it reduces to:

$$\sigma_{yy}(x-\mu_x)^2 \pm 2\sqrt{\sigma_{xx}\sigma_{yy}}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2 = 0$$

$$\Longleftrightarrow \quad \sqrt{\sigma_{yy}}(x-\mu_x) \pm \sqrt{\sigma_{xx}}(y-\mu_y) = 0$$

Thus the ellipse is *degenerated* to a straight line segment.
Next consider the following fact:

$$\left(\frac{\sigma_{xx}-\sigma_{yy}}{2}\right)^2 + \sigma_{xy}^2 \geq 0$$

On the other hand:

$$\left(\frac{\sigma_{xx}-\sigma_{yy}}{2}\right)^2 + \sigma_{xy}^2 = \left(\frac{\sigma_{xx}+\sigma_{yy}}{2}\right)^2 - \left(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2\right) \geq 0$$

$$\Longrightarrow \quad 0 \leq \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 \leq \left(\frac{\sigma_{xx}+\sigma_{yy}}{2}\right)^2$$

And the maximum is obtained for:

$$\left(\frac{\sigma_{xx}-\sigma_{yy}}{2}\right)^2 + \sigma_{xy}^2 = 0 \quad \Longleftrightarrow \quad \sigma_{xx} = \sigma_{yy} \quad \text{and} \quad \sigma_{xy} = 0$$

$$\Longleftrightarrow \quad \left(\frac{\sigma_{xx}+\sigma_{yy}}{2}\right)^2 = \sigma_{xx}^2 = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$$

Remember the Ellipse of Inertia:

$$\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2 = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$$

Which is now reduced to a circle:

$$\sigma_{xx}(x-\mu_x)^2 + \sigma_{xx}(y-\mu_y)^2 = \sigma_{xx}^2 \quad \Longleftrightarrow \quad (x-\mu_x)^2 + (y-\mu_y)^2 = \sigma_{xx}$$

Summarizing:
Determinant = minimum then Ellipse of Inertia = Straight Line Segment.
Determinant = maximum then Ellipse of Inertia = Circle.
Thus the determinant $\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$, eventually divided by the trace $(\sigma_{xx}+\sigma_{yy})/2$, is a $0 \leq$ measure ≤ 1 for the "roundness" of the ellipse.

Disclaimers

Anything free comes without referee :-(
My English may be better than your Dutch.