# Mathematical Prerequisites

A computer program has been written for the purpose of testing some - I hope so - new concepts in Pattern Recognition. The classical theory of moments has played a predominant role in the development of the code. This is exemplified by the Delphi unit **ogenblik.pas** , which is entirely devoted to it. As an application of the theory, a technique has been developed for the rendering of good old IBM punchcards. This document contains the mathematical theory of moments and some associated material.

## One-dimensional Moments

Consider a collection $X$ of arbitrary points $x_k$ in one-dimensional space. The members of this *points cloud* can be thought as coordinate positions on a straight line:

$$X = \{x_1, x_2, x_3, ..., x_k, ..., x_{N-1}, x_N\}$$

A quantity called weight or mass $m_k$ is associated with each of these points. The total weight or mass $M$ of the points is given by the sum of the partial weights $m_k$ :

$$M = \sum_{k=1}^{N} m_k = \sum_k m_k$$

It will be assumed in the sequel that the weights are always positive, meaning that they can be *normed.* Such normed weights $w_k$ are defined by:

$$w_k = \frac{m_k}{M} \quad \Longrightarrow \quad 0 \le w_k \le 1 \quad \text{and} \quad \sum_k w_k = 1$$

It is remarked that the weights $w_k$ can be interpreted as the components of a discrete probability distribution. Reason why we are tempted to conceive a certain spot, called center of mass, center of gravity, midpoint, middle or simply the *mean.* It is defined by:

$$\mu_x = \overline{x} = \sum_k w_k x_k$$

The midpoint takes a special position at the points cloud, since it is the weighted mean value of all positions of the points in the points cloud. It's easy to conceive a weighted mean value of other quantities, however. A most useful quantity is the so-called *second order moment*, which is also known as the *moment of inertia*, due to its applications in classical mechanics. Accordingly, the midpoint is also called a *first order moment.* The second order moment may also be called (the square of the) standard deviation or *spread*, due to the quite analogous quantity in Probability Theory:

$$\overline{x^2} = \sum_k w_k x_k^2$$

In addition to the above discrete quantities, there also exist *continuum versions* of the first and second order moments. The only difference is that the latter are defined by (definite) integrals instead of sums:

$$M = \int_a^b m(x)\,dx = \int m(x)\,dx \quad \text{and} \quad w(x) = \frac{m(x)}{M} \quad \Longrightarrow \quad \int w(x)\,dx = 1$$

$$\overline{x} = \int w(x)\,x\,dx \quad \text{and} \quad \overline{x^2} = \int w(x)\,x^2\,dx$$

It is clear from the outset, however, that such integrals are just limiting cases of discrete sums. Hence subsequent results will also be valid for the continuous version of the theory.
Second order moments may be defined with respect to a fixed, but otherwise arbitrary point $p$ in (1-D) space:

$$\sigma_{xx}(p) = \sum_k w_k(x_k - p)^2 \quad \text{or} \quad \sigma_{xx}(p) = \int w(x)\,(x-p)^2\,dx$$

The moment of inertia is interpreted as a mean of the squared distances of the points in the cloud with respect to a fixed point $p$. It will be shown now that there exists a preferrable origin, which is precisely the midpoint of the points distribution.

$$\sigma_{xx}(p) = \sum_k w_k(x_k - p)^2 = \sum_k w_k x_k^2 - 2p\sum_k w_k x_k + p^2 =$$

$$\overline{x^2} - 2p\overline{x} + p^2 = \left[\overline{x^2} - \overline{x}^2\right] + \left[\overline{x}^2 - 2p\overline{x} + p^2\right]$$

The first term between square brackets [ ] can be worked out as follows:

$$\left[\sum_k w_k x_k^2 - \left(\sum_k w_k x_k\right)^2\right] =$$

$$\sum_k w_k x_k^2 - 2\sum_k w_k \left(\sum_L w_L x_L\right) x_k + \sum_k w_k \left(\sum_L w_L x_L\right)^2 =$$

$$\sum_k w_k \left[x_k^2 - 2\left(\sum_L w_L x_L\right) x_k + \left(\sum_L w_L x_L\right)^2\right] =$$

$$\sum_k w_k \left[x_k - \left(\sum_L w_L x_L\right)\right]^2 = \sum_k w_k\,(x_k - \overline{x})^2$$

And the second term between square brackets [ ] is:

$$\left[\overline{x}^2 - 2p\overline{x} + p^2\right] = \left(\overline{x} - p\right)^2$$

Conclusion:

$$\sum_k w_k (x_k - p)^2 = \sum_k w_k \left(x_k - \overline{x}\right)^2 + \left(\overline{x} - p\right)^2$$

Then we see that the first term is positive, because it is a sum of (weighted) squares. But also the second term is a square and hence positive. The latter assumes a minimum if it is exactly zero, that is if: $p = \overline{x}$ . Formally:

$$\sum_k w_k (x_k - p)^2 = \text{minimum}(p) \quad \Longleftrightarrow \quad p = \overline{x} = \sum_k w_k x_k$$

The physical interpretation of the above is that a moment of inertia assumes a minimal value with respect to the origin if that origin is coincident with the center of mass. A moment of inertia with respect to an origin which is different from the center of mass can be expressed as the sum of two moments: one which expresses the moment of inertia with respect to the midpoint plus one which expresses the moment of inertia of the midpoint with respect to the origin. Unless explicitly stated otherwise, it will be assumed in the sequel that all moments of inertia are defined with respect to the midpoint $\mu_x$ or all (squares of the) spreads with respect to the mean. Then we can drop the dependence on $(p)$ in:

$$\sigma_{xx} = \sum_k w_k (x_k - \mu_x)^2$$

## Example: Line Segment

The equation of a line segment between $x_1$ and $x_2$ is:

$$x = x_1 + \xi(x_2 - x_1) \qquad \text{where:} \qquad 0 \le \xi \le 1 \quad \Longrightarrow \quad dx = (x_2 - x_1)d\xi$$

It is assumed that weights are uniformly distributed across this line segment: $w(\xi) = 1$. The midpoint of the line segment is then:

$$\overline{x} = \frac{\int_{x_1}^{x_2} x\,dx}{\int_{x_1}^{x_2} dx} = \frac{\int_0^1 \left[x_1 + (x_2 - x_1)\xi\right](x_2 - x_1)d\xi}{(x_2 - x_1)} =$$

$$x_1 + (x_2 - x_1)\int_0^1 \xi\,d\xi = x_1 + (x_2 - x_1)\frac{1}{2} = \frac{1}{2}(x_1 + x_2)$$

The second moments of inertia is:

$$\overline{x^2} = \int_0^1 x^2\,d\xi = x_1^2 + 2x_1(x_2 - x_1)\int_0^1 \xi\,d\xi + (x_2 - x_1)^2\int_0^1 \xi^2\,d\xi =$$

$$x_1^2 + 2x_1(x_2 - x_1)\frac{1}{2} + (x_2 - x_1)^2\frac{1}{3} = x_1^2 + x_1x_2 - x_1^2 + x_2^2/3 - 2x_1x_2/3 + x_1^2/3 =$$

$$\frac{1}{3}\left(x_1^2 + x_2^2 + x_1x_2\right) \quad \Longrightarrow$$

$$\overline{x^2} - \overline{x}^2 = \frac{1}{3}\left(x_1^2 + x_2^2 + x_1x_2\right) - \frac{1}{4}\left(x_1^2 + x_2^2 + 2x_1x_2\right) =$$

$$\frac{1}{12}\left(x_1^2 + x_2^2 - 2x_1x_2\right) = \frac{1}{12}\left(x_1 - x_2\right)^2$$

Summarizing:

$$\mu_x = \frac{1}{2}(x_1 + x_2) \quad \text{and} \quad \sigma_{xx} = \frac{1}{12}\left(x_2 - x_1\right)^2$$

## Two-dimensional Moments

Consider an arbitrary 2-D distribution of points $(x_k, y_k)$ in the plane. Aagain, a quantity called weight or mass $w_k$ is associated with each of these points. And again, we can define a spot, called the midpoint, center of mass or whatever name is to be preferred:

$$\sum_k w_k = 1$$

$$\mu_x = \overline{x} = \sum_k w_k x_k \quad \text{and} \quad \mu_y = \overline{y} = \sum_k w_k y_k$$

This is the discrete form. The continuous alternative is:

$$\iint w(x,y)\, dx\, dy = 1$$

$$\mu_x = \overline{x} = \iint w(x,y) x\, dx\, dy \quad \text{and} \quad \mu_y = \overline{y} = \iint w(x,y) y\, dx\, dy$$

Second order momenta, also called moments of inertia, are defined with respect to an origin $(p, q)$:

$$\sigma_{xx}(p) = \sum_k w_k (x_k - p)^2 \quad \text{and} \quad \sigma_{yy}(q) = \sum_k w_k (y_k - q)^2$$

$$\sigma_{xy}(p,q) = \sum_k w_k (x_k - p)(y_k - q)$$

The continuous form is:

$$\sigma_{xx}(p) = \iint w(x,y)(x-p)^2\, dx\, dy \quad \text{and} \quad \sigma_{yy}(q) = \iint w(x,y)(y-q)^2\, dx\, dy$$

$$\sigma_{xy}(p,q) = \iint w(x,y)(x-p)(y-q)\, dx\, dy$$

It has already been shown that, at least for $\sigma_{xx}$ ,there exists a preferrable origin, which is precisely the center of mass / geometric mean of the points distribution:

$$\sum_k w_k (x_k - p)^2 = \text{minimum}(p) \quad \Longleftrightarrow \quad p = \overline{x} = \sum_k w_k x_k$$

In very much the same way (method: what's in a name) we can prove for $\sigma_{yy}$ :

$$\sum_k w_k (y_k - q)^2 = \text{minimum}(q) \quad \Longleftrightarrow \quad q = \overline{y} = \sum_k w_k y_k$$

How about the "mixed" second order moment $\sigma_{xy}$ ?

$$\sigma_{xy}(\overline{x}, \overline{y}) = \sum_k w_k (x_k - \overline{x})(y_k - \overline{y}) = \sum_k w_k x_k y_k - \sum_k w_k x_k \overline{y} - \sum_k w_k y_k \overline{x} + \overline{x}\,\overline{y} =$$

$$\sum_k x_k y_k - \overline{x}\,\overline{y} - \overline{y}\,\overline{x} + \overline{x}\,\overline{y} \quad \Longrightarrow \quad \sigma_{xy} = \overline{xy} - \overline{x}\,\overline{y}$$

Again, unless explicitly stated otherwise, it will be assumed in the sequel that all moments of inertia are with respect to the midpoint $(\mu_x, \mu_y)$. Then we can drop $(p, q)$ in:

$$\sigma_{xx} = \sum_k w_k (x_k - \mu_x)^2 \quad \text{and} \quad \sigma_{yy} = \sum_k w_k (y_k - \mu_y)^2$$

$$\sigma_{xy} = \sum_k w_k (x_k - \mu_x)(y_k - \mu_y)$$

So far, it is less clear what kind of physical meaning should be attached to the quantity $\sigma_{xy}$ , which is known as a "cross correlation" in probability theory and statistics. Well, to be precise:

$$\rho = \frac{\sigma_{xy}}{\sqrt{\sigma_{xx}\sigma_{yy}}}$$

Where $\rho$ is the so-called *cross-correlation coefficient*. Suppose however, that we don't like $\sigma_{xy}$ at all and we only want to get rid of this term. How then could such a thing be accomplished ? It can certainly not be done by translation, since the origin of our coordinate system has already become fixed at the midpoint. But there is another possibility. It could be done by *rotating* the coordinate system in such a way that $\sigma'_{xy}$ becomes zero in the new ('primed') system. Let's give it a try. Start with:

$$\begin{cases} x' = cos(\theta)x + sin(\theta)y \\ y' = -sin(\theta)x + cos(\theta)y \end{cases}$$

Then:

$$\sigma'_{xy} = 0 \quad \Longleftrightarrow \quad \sum_k w_k x'_k y'_k =$$

$$\sum_k w_k \left[cos(\theta)x_k + sin(\theta)y_k\right]\left[-sin(\theta)x_k + cos(\theta)y_k\right] =$$

$$-cos(\theta)sin(\theta)\sum_k w_k x_k^2 + sin(\theta)cos(\theta)\sum_k w_k y_k^2$$

$$+\left[cos^2(\theta) - sin^2(\theta)\right]\sum_k w_k x_k y_k$$

Two goniometric formulas should me memorized here:

$$sin(2\theta) = 2sin(\theta)cos(\theta) \quad \text{and} \quad cos(2\theta) = cos^2(\theta) - sin^2(\theta)$$

Herewith:

$$\sum_k w_k x'_k y'_k = -\frac{1}{2} sin(2\theta)(\sigma_{xx} - \sigma_{yy}) + cos(2\theta)\sigma_{xy} = 0$$

Resulting in:

$$tan(2\theta) = \frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}} \qquad \text{for} \quad \sigma_{xx} \neq \sigma_{yy}$$

And as special cases:

$$sin(2\theta) = 0 \quad \Longrightarrow \quad \theta = k.\frac{\pi}{2} \quad k = 0, 1, 2, 3 \ldots$$

$$\text{for} \quad \sigma_{xx} \neq \sigma_{yy} \quad \text{and} \quad \sigma_{xy} = 0$$

Meaning that the situation where $\sigma_{xy} = 0$ is found back with every rotation of the coordinate system over 90 degrees.

$$cos(2\theta) = 0 \quad \Longrightarrow \quad \theta = \frac{\pi}{4} + k.\frac{\pi}{2} \quad k = 0, 1, 2, 3 \ldots$$

$$\text{for} \quad \sigma_{xx} = \sigma_{yy} \quad \text{and} \quad \sigma_{xy} \neq 0$$

Meaning that the situation where $\sigma_{xx} = \sigma_{yy}$ occurs every time $\theta$ is precisely in the middle between two angles where $\sigma_{xy} = 0$.
If both $\sigma_{xx} = \sigma_{yy}$ and $\sigma_{xy} = 0$ then the choice of the angle $\theta$ is arbitrary.
The other two moments of inertia, $\sigma'_{xx}$ and $\sigma'_{yy}$, are expressed into the angle $\theta$ as follows:

$$\sigma'_{xx} = \sum_k w_k (x'_k)^2 = \sum_k w_k \left[cos(\theta)x_k + sin(\theta)y_k\right]^2$$

$$= cos^2(\theta) \sum_k w_k x_k^2 + sin^2(\theta) \sum_k w_k y_k^2 + 2sin(\theta)cos(\theta) \sum_k w_k x_k y_k$$

$$\Longrightarrow \quad \sigma'_{xx} = cos^2(\theta)\sigma_{xx} + sin^2(\theta)\sigma_{yy} + 2sin(\theta)cos(\theta)\sigma_{xy}$$

And:

$$\sigma'_{yy} = \sum_k w_k (y'_k)^2 = \sum_k w_k \left[-sin(\theta)x_k + cos(\theta)y_k\right]^2$$

$$= sin^2(\theta) \sum_k w_k x_k^2 + cos^2(\theta) \sum_k w_k y_k^2 - 2sin(\theta)cos(\theta) \sum_k w_k x_k y_k$$

$$\Longrightarrow \quad \sigma'_{yy} = sin^2(\theta)\sigma_{xx} + cos^2(\theta)\sigma_{yy} - 2sin(\theta)cos(\theta)\sigma_{xy}$$

Working out the latter formula somewhat further:

$$\sigma'_{yy} = \left[1 - cos^2(\theta)\right]\sigma_{xx} + \left[1 - sin^2(\theta)\right]\sigma_{yy} - 2sin(\theta)cos(\theta)\sigma_{xy}$$

$$= \sigma_{xx} + \sigma_{yy} - \sigma'_{xx}$$

It is thus seen that the sum of the two "main" moments of inertia is entirely *invariant* for an orthogonal coordinate transformation:

$$\sigma'_{xx} + \sigma'_{yy} = \sigma_{xx} + \sigma_{yy}$$

We conclude that, indeed, the "unwanted" $\sigma_{xy}$ can be eliminated by a suitable rotation of the coordinate system, while the sum of the other "main" moments of inertia $(\sigma_{xx} + \sigma_{yy})$ remains independent of any such transformation. The question may be raised for what values of $\theta$ the transformed moments of inertia $\sigma'_{xx}$ and/or $\sigma'_{yy}$ attain an extreme value, a maximum or a minimum. In order to find out, derivatives to $\theta$ will be calculated. First we repeat:

$$\sigma'_{xx} = cos^2(\theta)\sigma_{xx} + sin^2(\theta)\sigma_{yy} + 2sin(\theta)cos(\theta)\sigma_{xy}$$

$$\sigma'_{yy} = sin^2(\theta)\sigma_{xx} + cos^2(\theta)\sigma_{yy} - 2sin(\theta)cos(\theta)\sigma_{xy}$$

Giving:

$$\frac{d}{d\theta}\sigma'_{xx} = -2sin(\theta)cos(\theta)\sigma_{xx} + 2cos(\theta)sin(\theta)\sigma_{yy} + 2cos^2(\theta)\sigma_{xy} - 2sin^2(\theta)\sigma_{xy}$$

$$= -sin(2\theta)(\sigma_{xx} - \sigma_{yy}) + cos(2\theta)2\sigma_{xy} = 0$$

$$\frac{d}{d\theta}\sigma'_{yy} = +2cos(\theta)sin(\theta)\sigma_{xx} - 2sin(\theta)cos(\theta)\sigma_{yy} - 2cos^2(\theta)\sigma_{xy} + 2sin^2(\theta)\sigma_{xy}$$

$$= +sin(2\theta)(\sigma_{xx} - \sigma_{yy}) - cos(2\theta)2\sigma_{xy} = 0$$

It is seen that exactly the same equations are obtained as with $\sigma'_{xy} = 0$. Meaning that $\sigma'_{xx}$ and $\sigma'_{yy}$ attain their extreme values, both at the same time, when and only when $\sigma'_{xy} = 0$. Opposite signs indicate that one of the two extremes, $\sigma'_{xx}$ or $\sigma'_{yy}$, must be a minimum while the other must be a maximum.
Alternative viewpoints may be obtained by just reversing the whole story. Take the 'primed' coordinate system for granted. And transform back to 'unprimed' coordinates. In order to accomplish this, it is not necessary to solve the above equations for the unprimed moments of inertia. Instead, simply reverse the angle of rotation and you're done:

$$\sigma_{xx} = cos^2(-\theta)\sigma'_{xx} + sin^2(-\theta)\sigma'_{yy} + 2sin(-\theta)cos(-\theta)\sigma'_{xy}$$

$$\sigma_{yy} = sin^2(-\theta)\sigma'_{xx} + cos^2(-\theta)\sigma'_{yy} - 2sin(-\theta)cos(-\theta)\sigma'_{xy}$$

$$\sigma_{xy} = -\frac{1}{2}sin(-2\theta)(\sigma'_{xx} - \sigma'_{yy}) + cos(-2\theta)\sigma'_{xy}$$

Remember, however, that the cross correllation moment is zero by definition in the primed system. At last, consider the primed system as the standard one. Herewith it is expressed that the coordinate system where the cross correllation moment is zero is to be considered in the sequel as the *preferred* system of coordinates. The x- and y-axis, associated with the preferred system, are known in Physics as the *main axes of inertia*. They are attached to the points cloud, as body fitted coordinates so to speak. Any other system is now the result of a rotation of the main axes of inertia, over a certain angle $\theta$ . And the transformed moments of inertia can always be derived from the main moments of inertia:

$$\sigma_{xx} = cos^2(\theta)\sigma'_{xx} + sin^2(\theta)\sigma'_{yy} \quad \text{and} \quad \sigma_{yy} = sin^2(\theta)\sigma'_{xx} + cos^2(\theta)\sigma'_{yy}$$

$$\sigma_{xy} = \frac{1}{2}sin(2\theta)(\sigma'_{xx} - \sigma'_{yy})$$

It is seen that $\sigma_{xy} = 0$ , for $\theta = k.\pi/2 \quad k = 1, 2, \ldots$ . The same angle values cause $\sigma_{xx}$ to become equal to $\sigma'_{xx}$, for $\theta = k.\pi$ , or equal to $\sigma'_{yy}$, for $\theta = \pi/2+k.\pi$. And the same angle values cause $\sigma_{yy}$ to become equal to $\sigma'_{yy}$, for $\theta = k.\pi$ , or equal to $\sigma'_{xx}$, for $\theta = \pi/2 + k.\pi$. It all means that $\sigma'_{xx}$ and $\sigma'_{yy}$ exchange roles with every increase of the angle $\theta$ with $90^o$ , while $\sigma_{xy} = 0$ at the same time.
On the other hand, $\sigma_{xx}$ and $\sigma_{yy}$ become equal for $\theta = \pi/4 + k.\pi/2$ . Then the cross correllation moment $\sigma_{xy}$ attains a minimum (negative) or maximum (positive) value of $\pm(\sigma'_{xx} - \sigma'_{yy})/4$ .
If $\sigma'_{xx} = \sigma'_{yy}$, that is when the main moments of inertia are equal to each other, then also the transformed main moments of inertia always will be equal to each other and the cross correllation moment $\sigma_{xy}$ will always be zero. This special case is often induced by symmetry.
Last but not least. The fact that the above formulas are entirely insensitive to an increase of the angle $\theta$ with $180^o$ also means that it is impossible to detect the *orientation* of the main axes coordinate system, with help of moments up to the second order alone. To that end, moments of at least order tree are needed.

### Example: Quarter of a Circle

As an example, consider the area which is delimited by:

$$x^2 + y^2 \leq R^2 \quad \text{and} \quad x \geq 0 \quad \text{and} \quad y \geq 0$$

The area of this quarter of a circle is simply given by:

$$A = \frac{1}{4}\pi R^2$$

The first order moments are calculated as follows:

$$\overline{x} = \frac{\iint x\, dxdy}{\iint dxdy} = \frac{\int_0^R \int_0^{\pi/2} r.cos(\phi)\, r.dr.d\phi}{\pi.R^2/4} = \frac{4}{\pi R^2} \int_0^R r^2\, dr \int_0^{\pi/2} cos(\phi)\, d\phi$$

$$= \frac{4}{\pi R^2} \left[ \frac{1}{3} r^3 \right]_0^R [\, sin(\phi)\, ]_0^{\pi/2} = \frac{4R}{3\pi} \approx 0.42441\, R$$

Likewise:

$$\overline{y} = \frac{\iint y\, dxdy}{\iint dxdy} = \frac{4}{\pi R^2} \left[ \frac{1}{3} r^3 \right]_0^R [\, -cos(\phi)\, ]_0^{\pi/2} = \frac{4R}{3\pi}$$

The second order moments are calculated as follows:

$$\overline{x^2} = \frac{\iint x^2\, dxdy}{\iint dxdy} = \frac{\int_0^R \int_0^{\pi/2} [r.cos(\phi)]^2\, r.dr.d\phi}{\pi.R^2/4} = \frac{4}{\pi R^2} \int_0^R r^3\, dr \int_0^{\pi/2} cos^2(\phi)\, d\phi$$

Where:

$$cos(2\phi) = 2.cos^2(\phi) - 1 \quad \Longrightarrow \quad cos^2(\phi) = \frac{cos(2\phi) + 1}{2} \quad \Longrightarrow$$

$$\int_0^{\pi/2} cos^2(\phi)\, d\phi = \frac{1}{4} \int_0^{\pi/2} cos(2\phi)\, d(2\phi) \; + \; \frac{1}{2} \int_0^{\pi/2} d\phi = \frac{1}{4} [\, sin(\theta)\, ]_0^{\pi} + \frac{1}{2}(\pi/2)$$

$$\Longrightarrow \quad \overline{x^2} = \frac{4}{\pi R^2} \left[ \frac{r^4}{4} \right]_0^R . \frac{\pi}{4} = \frac{R^2}{4}$$

In very much the same way we would find:

$$\overline{y^2} = \frac{R^2}{4}$$

This result can also be predicted, however, by considerations of symmetry. The whole figure is symmetric around the line $y = x$, namely, and so any result for $x$ can also be used for $y$ . The cross moment of inertia is:

$$\overline{xy} = \frac{4}{\pi R^2} \int_0^R r^2.r.dr \int_0^{\pi/2} cos(\phi) sin(\phi)\, d\phi = \frac{4}{\pi R^2} \left[ \frac{r^4}{4} \right]_0^R \frac{1}{2} \left[ sin^2(\phi) \right]_0^{\pi/2}$$

We find that it has only part of the magnitude of the two other second order moments:

$$\overline{xy} = \frac{R^2}{2\pi} < \frac{R^2}{4}$$

However, all second order moments must be calculated with respect to the center of mass:

$$\overline{x^2} - \overline{x}^2 = \overline{y^2} - \overline{y}^2 = \frac{R^2}{4} - \left(\frac{4R}{3\pi}\right)^2 \quad \text{and} \quad \overline{xy} - \overline{x}\overline{y} = \frac{R^2}{2\pi} - \left(\frac{4R}{3\pi}\right)^2$$

Recall the formulas:

$$\sigma'_{xx} = cos^2(\theta)\sigma_{xx} + sin^2(\theta)\sigma_{yy} + 2sin(\theta)cos(\theta)\sigma_{xy}$$

$$\sigma'_{yy} = sin^2(\theta)\sigma_{xx} + cos^2(\theta)\sigma_{yy} - 2sin(\theta)cos(\theta)\sigma_{xy}$$

In our case, we have typically the situation where the angle of rotation must be $45^o$ ($\theta = \pi/4$), in order to arrive at the main moments of inertia:

$$\sigma'_{xx} = \frac{1}{2}\sigma_{xx} + \frac{1}{2}\sigma_{yy} + 2\frac{1}{2}\sigma_{xy}$$

$$\sigma'_{yy} = \frac{1}{2}\sigma_{xx} + \frac{1}{2}\sigma_{yy} - 2\frac{1}{2}\sigma_{xy}$$

Substitute the above values:

$$\sigma'_{xx} = \frac{R^2}{4} - \left(\frac{4R}{3\pi}\right)^2 + \left[\frac{R^2}{2\pi} - \left(\frac{4R}{3\pi}\right)^2\right] \approx 0.04890\, R^2$$

$$\sigma'_{yy} = \frac{R^2}{4} - \left(\frac{4R}{3\pi}\right)^2 - \left[\frac{R^2}{2\pi} - \left(\frac{4R}{3\pi}\right)^2\right] \approx 0.09085\, R^2$$

## Least Squares Method

Suppose we have a collection of points in the plane and we want to draw a straight line through these points, in such a way that the line is a "best fit" to them. An equation for the straight line can always be set up as follows:

$$cos(\theta)(x-p)+sin(\theta)(y-q)=0$$

Here $\theta$ is the angle of the line's normal with the x-axis and $(p,q)$ is an arbitrary point on the line. The distance of an arbitrary point $\vec{r}=(x_k,y_k)$ to the line is given by the length of the projection of the point's vector $\vec{r}$ onto the normal $\vec{n}$ of the line. The latter is given by $\vec{n}=(cos(\theta),sin(\theta))$. Hence the length of the projection is:

$$\left|\frac{(\vec{r}\cdot\vec{n})}{(\vec{n}\cdot\vec{n})}\vec{n}\right| = |cos(\theta)(x_k-p)+sin(\theta)(y_k-q)|$$

For the straight line to be a "best fit", it will be required that the sum of the weighted squares of all distances shall be a minimum:

$$\sum_k w_k\left[cos(\theta)(x_k-p)+sin(\theta)(y_k-q)\right]^2 = \text{minimum}(p,q,\theta)$$

Working out a bit:

$$cos^2(\theta)\sum_k w_k(x_k-p)^2+sin^2(\theta)\sum_k w_k(y_k-q)^2+$$

$$2sin(\theta)cos(\theta)\sum_k w_k(x_k-p)(y_k-q) = \text{minimum}(p,q,\theta)$$

Let us solve just one part of the puzzle, namely: how the points $(p,q)$ must be selected in such a way that a minimum may be reached with respect to this choice. For certain parts of the above expression this would mean that:

$$\sum_k w_k(x_k-p)^2 = \text{minimum} \quad \text{and} \quad \sum_k w_k(y_k-q)^2 = \text{minimum}$$

We have already seen that minimal values are reached if second order momenta are described with respect to the midpoint of the points cloud as their origin. In our case:

$$p=\sum_k w_k x_k \quad \text{and} \quad q=\sum_k w_k y_k$$

Define second order momenta with respect to the midpoint as usual:

$$\sigma_{xx}=\sum_k w_k(x_k-p)^2 \quad \text{and} \quad \sigma_{yy}=\sum_k w_k(y_k-q)^2$$

$$\sigma_{xy} = \sum_k w_k(x_k - p)(y_k - q)$$

We can concentrate now on minimalization with respect to the angle $\theta$:

$$cos^2(\theta)\sigma_{xx} + sin^2(\theta)\sigma_{yy} + 2sin(\theta)cos(\theta)\sigma_{xy} = \text{minimum}(\theta)$$

Extreme values may be found by differentiation to the independent variable:

$$-2sin(\theta)cos(\theta)\sigma_{xx} + 2cos(\theta)sin(\theta)\sigma_{yy} + 2cos^2(\theta)\sigma_{xy} - 2sin^2(\theta)\sigma_{xy} = 0$$

Which leads to the familiar equation:

$$sin(2\theta)(\sigma_{xx} - \sigma_{yy}) - 2\,cos(2\theta)\sigma_{xy} = 0$$

The above expression is also recognized as the one which caused the "cross correlation moment" $\sigma_{xy}$ to become zero in a rotated coordinate system. It is an expression in *two times* the angle of the straight line with the x-axis. The meaning of it being that, if $\theta$ is a solution, then also $\theta + \pi/2$ must be a solution. Meaning in turn that besides the straight line itself also *the line perpendicular to it* is a solution. However, this is quite sensible because only *extrema* are found by differentiation and putting the outcome to zero. Indeed, one finds a minimum for one value of the angle $\theta$ and a maximum for the perpendicular angle. The wanted line (minimum) is a best fit to the points and the (unwanted) perpendicular line is a worst fit to the points. Both lines go through the midpoint or gravitational center of the points cloud.
It must be concluded herefrom that the two straight lines through the midpoint of the cloud, the one with the best fit as well as the one with the worst fit, together form an orthogonal coordinate system which is the same as the standard coordinate system, consisting of the *main* axes of inertia of the points cloud.

## Schwarz inequality

Inclining towards Linear Algebra, but with the theory of point clouds in mind, we will redefine the inner product of two vectors as:

$$(\vec{a} \cdot \vec{b}) := \sum_k w_k a_k b_k$$

It can be shown easily that such an inner product obeys all of the usual rules:

$$\begin{aligned}(\vec{a} \cdot \vec{b}) &= (\vec{b} \cdot \vec{a}) \\ (\vec{a} \cdot (\vec{b} + \vec{c})) &= (\vec{a} \cdot \vec{b}) + (\vec{a} \cdot \vec{c}) \\ (\vec{a} \cdot \lambda\vec{b}) &= \lambda(\vec{a} \cdot \vec{b}) \\ (\vec{a} \cdot \vec{a}) &\geq 0\end{aligned}$$

For the last rule to be obeyed, it is necessary that masses $w_k$ be positive (or zero). Schwartz inequality can now be conjectured for this inner product:

$$(\vec{a} \cdot \vec{b})^2 \leq (\vec{a} \cdot \vec{a})(\vec{b} \cdot \vec{b})$$

Proof:

$$(\lambda\vec{a} - \vec{b} \cdot \lambda\vec{a} - \vec{b})^2 \geq 0 \quad \Longrightarrow$$

$$\lambda^2(\vec{a} \cdot \vec{a}) - 2\lambda(\vec{a} \cdot \vec{b}) + (\vec{b} \cdot \vec{b}) \geq 0$$

This is a quadratic inequality in $\lambda$. In order for this inequality to hold, its discriminant must be negative or zero:

$$4(\vec{a} \cdot \vec{b})^2 - 4(\vec{a} \cdot \vec{a})(\vec{b} \cdot \vec{b}) \leq 0 \quad \Longrightarrow$$

$$(\vec{a} \cdot \vec{b})^2 \leq (\vec{a} \cdot \vec{a})(\vec{b} \cdot \vec{b})$$

Which completes the proof. Returning now to our original problem, define for example:

$$\vec{a} = (x_1, x_2, ..., x_k, ..., x_N) \quad \text{and} \quad \vec{b} = (1, 1, ..., 1, ..., 1)$$

Then:

$$(\vec{a} \cdot \vec{b}) = \sum_k w_k x_k \quad \text{and} \quad (\vec{a} \cdot \vec{a}) = \sum_k w_k x_k^2 \quad \text{and} \quad (\vec{b} \cdot \vec{b}) = \sum_k w_k = 1$$

Therefore:

$$\left(\sum_k w_k x_k\right)^2 \leq \left(\sum_k w_k x_k^2\right) \quad \Longrightarrow \quad \sum_k w_k x_k^2 - \left(\sum_k w_k x_k\right)^2 \geq 0$$

This result is equivalent with our previous finding that:

$$\sum_k w_k (x_k - \overline{x})^2 \geq 0$$

New results are obtained when applying Schwarz inequality to the quantity known formerly as cross correllation moment. For the sake of simplicity, the midpoint is set as the origin $(0, 0)$ of the coordinate system. Now define:

$$\vec{a} = (x_1, x_2, ..., x_k, ..., x_N) \quad \text{and} \quad \vec{b} = (y_1, y_2, ..., y_k, ..., y_N)$$

Then:

$$(\vec{a} \cdot \vec{b}) = \sum_k w_k x_k y_k = \sigma_{xy}$$

$$(\vec{a} \cdot \vec{a}) = \sum_k w_k x_k^2 = \sigma_{xx} \quad \text{and} \quad (\vec{b} \cdot \vec{b}) = \sum_k w_k y_k^2 = \sigma_{yy}$$

Therefore, according to Schwarz inequality, the following relationship must hold:

$$\sigma_{xy}^2 \leq \sigma_{xx}\sigma_{yy} \quad \text{or} \quad \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 \geq 0$$

As a consequence, the following quantity is bound like a (co)sinus-function:

$$\rho := \frac{\sigma_{xy}}{\sqrt{\sigma_{xx}\sigma_{yy}}} \qquad \text{where} \qquad 0 \leq |\rho| \leq 1$$

The thus defined quantity $\rho$ has already been mentioned as the cross-correlation, in a narrower sense.

## Tensor of Inertia

In two dimensional space, the first order moments can be conceived as the two components of a vector:

$$\vec{\mu} = \begin{bmatrix} \sum_k w_k x_k \\ \sum_k w_k y_k \end{bmatrix}$$

Likewise, the second order moments can be conceived as the three components of a *symmetric matrix*, the so-called inertial tensor:

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

At a more sophisticated level, the problem of finding the main axes of inertia can then be approached via Linear Algebra, especially the theory of eigenvalues and eigenvectors. It's a matter of routine to show that the expressions found for the transformed moments of inertia are equivalent with the following: find the orthogonal transformation (i.e. rotation over an angle $\theta$) which reduces the tensor (matrix) of inertia to its diagonal form:

$$\begin{bmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{bmatrix} \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix} \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} = \begin{bmatrix} \sigma'_{xx} & 0 \\ 0 & \sigma'_{yy} \end{bmatrix}$$

This, in turn, is equivalent with finding the eigenvalues $\lambda$ and the eigenvectors $(\kappa_x, \kappa_y)$ of the inertial matrix:

$$\begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix} \begin{bmatrix} \kappa_x \\ \kappa_y \end{bmatrix} = \lambda \begin{bmatrix} \kappa_x \\ \kappa_y \end{bmatrix}$$

The corresponding characteristic equations are:

$$\begin{vmatrix} \sigma_{xx} - \lambda & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} - \lambda \end{vmatrix} = 0 \quad \Longleftrightarrow$$

$$(\sigma_{xx} - \lambda)(\sigma_{yy} - \lambda) - \sigma_{xy}^2 = 0 \quad \Longleftrightarrow \quad \lambda^2 - (\sigma_{xx} + \sigma_{yy})\lambda + (\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2) = 0$$

Define trace $Sp$ and determinant $Det$ by:

$$Sp := \sigma_{xx} + \sigma_{yy} \quad \text{and} \quad Det := \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$$

We have seen that the trace $(\sigma_{xx} + \sigma_{yy})$ is invariant for a rotation of the coordinate system. It is remarked that the determinant $(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)$ is also *invariant* for such a transformation. The latter can be accepted as a well known fact of Linear Algebra, or proved with help of elementary, though somewhat laborious goniometry.

Having established trace and determinant, the characteristic equation of the eigenvalue problem can be written as:

$$\lambda^2 - (Sp)\lambda + Det = 0$$

Most of the time, a quadratic equation has two solutions. The greatest of the two solutions will be called $\lambda_1$ and the smallest one $\lambda_2$. Sum and product of the solutions are found immediately:

$$\lambda_1 + \lambda_2 = Sp \quad \text{and} \quad \lambda_1 . \lambda_2 = Det$$

Write the equation as:

$$\lambda^2 - 2(Sp/2)\lambda + (Sp/2)^2 = (Sp/2)^2 - Det \quad \Longleftrightarrow$$

$$[\lambda - (Sp/2)]^2 = (Sp/2)^2 - Det \quad \Longleftrightarrow \quad \lambda - (Sp/2) = \pm\sqrt{(Sp/2)^2 - Det}$$

$$\Longleftrightarrow \quad \lambda = (Sp/2) \pm \sqrt{(Sp/2)^2 - Det}$$

Herewith we find the solutions:

$$\lambda_1 = (Sp/2) + \sqrt{(Sp/2)^2 - Det}$$

$$\lambda_2 = (Sp/2) - \sqrt{(Sp/2)^2 - Det} = Det/\lambda_1$$

Provided that de discriminant is positive, indeed:

$$(Sp/2)^2 - Det = \frac{1}{4}(\sigma_{xx} + \sigma_{yy})^2 - (\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2) =$$

$$\frac{1}{4}\sigma_{xx}^2 + \frac{1}{4}\sigma_{yy}^2 + \frac{1}{2}\sigma_{xx}\sigma_{yy} - \sigma_{xx}\sigma_{yy} + \sigma_{xy}^2 = \frac{1}{4}\sigma_{xx}^2 + \frac{1}{4}\sigma_{yy}^2 - \frac{1}{2}\sigma_{xx}\sigma_{yy} + \sigma_{xy}^2 =$$

$$= \frac{1}{4}(\sigma_{xx} - \sigma_{yy})^2 + \sigma_{xy}^2 > 0$$

Herewith the eigenvalues can also be expressed as:

$$\lambda_{12} = \frac{1}{2}(\sigma_{xx} + \sigma_{yy}) \pm \frac{1}{2}(\sigma_{xx} - \sigma_{yy})\sqrt{1 + \left(\frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}}\right)^2}$$

Where the expression between parentheses ( ) is recognized as:

$$\frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}} = tan(2\theta)$$

Calculation of the accompanying eigenvectors may be a tricky business, unless proper measures are taken. Reconsider the formula which expresses the angle over which the coordinate system must be rotated in order to force the cross correllation moment to be zero:

$$tan(2\theta) = \frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}} \qquad \text{for} \quad \sigma_{xx} \neq \sigma_{yy}$$

It is quite useful to distinguish the special cases $\sigma_{xy} = 0$ and $\sigma_{xx} = \sigma_{yy}$ in the first place. But then the formula readily gives an angle $\theta$ :

$$\theta = \frac{1}{2} arctan \left( \frac{2\sigma_{xy}}{\sigma_{xx} - \sigma_{yy}} \right)$$

Resulting in two possible eigenvectors, because the angle $\theta$ is ambiguous up to a multiple of $90^o$ :

$$\begin{bmatrix} cos(\theta) \\ sin(\theta) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -sin(\theta) \\ cos(\theta) \end{bmatrix}$$

It is necessary to decide afterwards which of the two eigenvectors is the one belonging to the chosen eigenvalue. To that end, multiply both vectors with the (singular) eigenvalue matrix:

$$\begin{bmatrix} \sigma_{xx} - \lambda_1 & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} - \lambda_1 \end{bmatrix} \begin{bmatrix} cos(\theta) \\ sin(\theta) \end{bmatrix} \begin{bmatrix} -sin(\theta) \\ cos(\theta) \end{bmatrix}$$

The outcome with the smallest vector-length (idealiter = zero) corresponds with the eigenvector to be selected.

# Triangle Algebra

Let's consider the simplest non-trivial finite element shape in two dimensions: the linear triangle. Function behaviour is approximated inside such a triangle by a *linear* interpolation between the function values at the vertices, also called: nodal points. Let $T$ be such a function, and $x, y$ coordinates, then:

$$T = A.x + B.y + C$$

Where the constants A, B, C are yet to be determined.

Substitute $x = x_k$ and $y = y_k$ with $k = 1, 2, 3$:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} C \\ A \\ B \end{bmatrix}$$

The first of these equations can already be used to eliminate the constant $C$, once and forever:

$$T_1 = A.x_1 + B.y_1 + C$$

Resulting in:

$$T - T_1 = A.(x - x_1) + B.(y - y_1)$$

Hence the constants $A$ and $B$ are determined by:

$$\begin{aligned} T_2 - T_1 = A.(x_2 - x_1) + B.(y_2 - y_1) \\ T_3 - T_1 = A.(x_3 - x_1) + B.(y_3 - y_1) \end{aligned}$$

Two equations with two unknowns. The solution is found by straightforward elimination, or by applying Cramer's rule:

$$\begin{aligned} A = [(y_3 - y_1).(T_2 - T_1) - (y_2 - y_1).(T_3 - T_1)]/\Delta \\ B = [(x_2 - x_1).(T_3 - T_1) - (x_3 - x_1).(T_2 - T_1)]/\Delta \end{aligned}$$

There are several forms of the determinant $\Delta$, which should be memorized when it is appropriate:

$$\Delta = (x_2 - x_1).(y_3 - y_1) - (x_3 - x_1).(y_2 - y_1)$$
$$\Delta = 2 \times \text{ area of triangle}$$
$$\Delta = x_1.y_2 + x_2.y_3 + x_3.y_1 - y_1.x_2 - y_2.x_3 - y_3.x_1$$
$$\Delta = x_1.(y_2 - y_3) + x_2.(y_3 - y_1) + x_3.(y_1 - y_2)$$
$$\Delta = y_1.(x_3 - x_2) + y_2.(x_1 - x_3) + y_3.(x_2 - x_1)$$
$$\Delta = \left\| \begin{array}{ccc} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{array} \right\|$$

Anyway, it is concluded that:

$$T - T_1 = \xi.(T_2 - T_1) + \eta.(T_3 - T_1)$$

Where:

$$\xi = [(y_3 - y_1).(x - x_1) - (x_3 - x_1).(y - y_1)]/\Delta$$
$$\eta = [(x_2 - x_1).(y - y_1) - (y_2 - y_1).(x - x_1)]/\Delta$$

Or:

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} +(y_3 - y_1) & -(x_3 - x_1) \\ -(y_2 - y_1) & +(x_2 - x_1) \end{bmatrix} / \Delta \begin{bmatrix} x - x_1 \\ y - y_1 \end{bmatrix}$$

The inverse of the following problem is recognized herein:

$$\begin{bmatrix} x - x_1 \\ y - y_1 \end{bmatrix} = \begin{bmatrix} (x_2 - x_1) & (x_3 - x_1) \\ (y_2 - y_1) & (y_3 - y_1) \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}$$

Or:

$$x - x_1 = \xi.(x_2 - x_1) + \eta.(x_3 - x_1)$$
$$y - y_1 = \xi.(y_2 - y_1) + \eta.(y_3 - y_1)$$

But also:

$$T - T_1 = \xi.(T_2 - T_1) + \eta.(T_3 - T_1)$$

Therefore the *same* expression holds for the function $T$ as well as for the coordinates $x$ and $y$ . This is precisely what people mean by an *isoparametric* ("same parameters") transformation, a terminology which is quite common in Finite Element contexts. Now recall the formulas which express $\xi$ and $\eta$ into $x$ and $y$ :

$$\xi = [(y_3 - y_1).(x - x_1) - (x_3 - x_1).(y - y_1)]/\Delta$$
$$\eta = [(x_2 - x_1).(y - y_1) - (y_2 - y_1).(x - x_1)]/\Delta$$

Thus $\xi$ can be interpreted as: area of the sub-triangle spanned by the vectors $(x-x_1, y-y_1)$ and $(x_3-x_1, y_3-y_1)$ divided by the whole triangle area. And $\eta$ can be interpreted as: area of the sub-triangle spanned by the vectors $(x-x_1, y-y_1)$

and $(x_2 - x_1, y_2 - y_1)$ divided by the whole triangle area. This is the reason why $\xi$ and $\eta$ are sometimes called *area-coordinates*; see the above figure, where (two times) the area of the triangle as a whole is denoted as $\Delta$. There exist even *three* of these coordinates in literature. But the third area-coordinate is, of course, dependent on the other two, being equal to $(1 - \xi - \eta)$. Instead of area-coordinates, we prefer to talk about *local coordinates* $\xi$ and $\eta$ of an element, in contrast to the *global coordinates* $x$ and $y$. It is possible that local coordinates coincide with the global coordinates. A triangle for which such is the case is called a *parent element*. The portrait of the parent triangle is also depicted in the above figure: it is rectangular, and two sides of it are equal.
Let's reconsider the expression:

$$T - T_1 = \xi.(T_2 - T_1) + \eta.(T_3 - T_1)$$

Partial differentiation to $\xi$ and $\eta$ gives:

$$\partial T/\partial\xi = T_2 - T_1 \quad ; \quad \partial T/\partial\eta = T_3 - T_1$$

Therefore, with node (1) as the origin, hence $T(0) = T_1$:

$$T = T(0) + \xi\frac{\partial T}{\partial\xi} + \eta\frac{\partial T}{\partial\eta}$$

This is part of a Taylor series expansion around node (1). Such Taylor series expansions are quite common in Finite Difference analysis. Now rewrite as follows:

$$T = (1 - \xi - \eta).T_1 + \xi.T_2 + \eta.T_3$$

Here the functions $(1 - \xi - \eta)$, $\xi$, $\eta$ are called the *shape functions* of the Finite Element. Shape functions $N_k$ have the property that they are unity in one of the nodes (k), and zero in all other nodes. In our case:

$$N_1 = 1 - \xi - \eta \quad ; \quad N_2 = \xi \quad ; \quad N_3 = \eta$$

So we have two representations, which are almost trivially equivalent:

$$T = T_1 + \xi.(T_2 - T_1) + \eta.(T_3 - T_1) \qquad \text{: Finite Difference like}$$
$$T = (1 - \xi - \eta).T_1 + \xi.T_2 + \eta.T_3 \qquad \text{: Finite Element like}$$

What kind of terms can be discretized at the domain of a linear triangle? In the first place, the function $T(x, y)$ itself, of course. But one may also try the first order partial derivatives $\partial T/\partial x$ , $\partial T/\partial y$. We find:

$$\partial T/\partial x = A = [(y_3 - y_1).(T_2 - T_1) - (y_2 - y_1).(T_3 - T_1)]/\Delta$$
$$\partial T/\partial y = B = [(x_2 - x_1).(T_3 - T_1) - (x_3 - x_1).(T_2 - T_1)]/\Delta$$

By collecting terms belonging to the same $T_k$, this can also be written as:

$$\Delta \begin{bmatrix} \partial T/\partial x \\ \partial T/\partial y \end{bmatrix} = \begin{bmatrix} +(y_2 - y_3) & +(y_3 - y_1) & +(y_1 - y_2) \\ -(x_2 - x_3) & -(x_3 - x_1) & -(x_1 - x_2) \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}$$

Or, in operator form:

$$\begin{bmatrix} \partial/\partial x \\ \partial/\partial y \end{bmatrix} = \begin{bmatrix} y_2 - y_3 & y_3 - y_1 & y_1 - y_2 \\ x_3 - x_2 & x_1 - x_3 & x_2 - x_1 \end{bmatrix} /\Delta$$

The right hand side will be called a *Differentiation Matrix* in subsequent work. Thus the gradient operator at any linear triangle is represented by a $2 \times 3$ differentiation matrix.

## Triangle Integrals

Our goal is to calculate some (first, second, third order) moments of an arbitrary triangle. To be mathematically precise:

$$\mu_x = \iint x\,dxdy \quad \text{and} \quad \mu_y = \iint y\,dxdy$$

$$\sigma_{xx} = \iint x^2\,dxdy \quad \text{and} \quad \sigma_{yy} = \iint y^2\,dxdy \quad \text{and} \quad \sigma_{xy} = \iint xy\,dxdy$$

Or even:

$$M_{xxx} = \iint x^3\,dxdy \quad \text{and} \quad M_{xxy} = \iint x^2 y\,dxdy$$

$$M_{xyy} = \iint xy^2\,dxdy \quad \text{and} \quad M_{yyy} = \iint y^3\,dxdy$$

Following the theory in the previous paragraph, global coordinates $(x, y)$ can be expressed in their local counterparts $(\xi, \eta)$ :

$$x - x_0 = \xi.(x_1 - x_0) + \eta.(x_2 - x_0)$$
$$y - y_0 = \xi.(y_1 - y_0) + \eta.(y_2 - y_0)$$

It makes no difference for the outcome of the integrals if a more handsome choice for the coordinate system is to be preferred. Therefore, let one of the vertices of the triangle, say $(x_0, y_0)$, be selected as the origin of our global coordinate system. Then:

$$x = \xi.x_1 + \eta.x_2 \quad \text{and} \quad y = \xi.y_1 + \eta.y_2$$

And the Jacobian $\Delta$ of this transformation is involved with:

$$dx.dy = (x_1 y_2 - x_2 y_1)\,d\xi d\eta = \Delta\,d\xi d\eta$$

Limited use will be made of Newton's binomial formula:

$$(a+b)^n = \sum_{k=0}^{n} C(n,k)\,a^k b^{n-k} = \sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!}\,a^k b^{n-k}$$

The formulas for any moment of a triangle take the following general form:

$$\iint x^m y^n\,dxdy = \iint (\xi.x_1 + \eta.x_2)^m (\xi.y_1 + \eta.y_2)^n\,\Delta\,d\xi d\eta =$$

$$\iint \sum_{i=0}^{m} C(m,i)\,\xi^i x_1^i \eta^{m-i} x_2^{m-i} \sum_{j=0}^{n} C(n,j)\,\xi^j y_1^j \eta^{n-j} y_2^{n-j}\,\Delta\,d\xi d\eta =$$

$$\Delta \sum_{i=0}^{m} \sum_{j=0}^{n} C(m,i)\,x_1^i x_2^{m-i} C(n,j)\,y_1^j y_2^{n-j} \iint \xi^{i+j} \eta^{m+n-i-j}\,d\xi d\eta$$

The above formula - being far too complicated - will not be used in the sequel. It turns out, however, that we have to calculate integrals like:

$$F(m,n) = \iint \xi^m \eta^n \, d\xi d\eta$$

Herewith the integration is carried out over a rectangular equilateral triangle, with local coordinates $\xi$ and $\eta$, where $0 \le \xi \le 1$ and $0 \le \eta \le (1-\xi)$ . Working out a few steps:

$$F(m,n) = \iint \xi^m \eta^n \, d\xi d\eta = \int_0^1 \xi^m \left[ \int_0^{1-\xi} \eta^n d\eta \right] d\xi$$

$$= \int_0^1 \xi^m \left[ \frac{(1-\xi)^{n+1}}{n+1} \right] d\xi = \int_0^1 \frac{(1-\xi)^{n+1}}{n+1} d\left( \frac{\xi^{m+1}}{m+1} \right)$$

$$= \left[ \frac{(1-\xi)^{n+1}}{n+1} \frac{\xi^{m+1}}{m+1} \right]_0^1 - \int_0^1 \frac{\xi^{m+1}}{m+1} d\left( \frac{(1-\xi)^{n+1}}{n+1} \right) = 0 + \int_0^1 \frac{\xi^{m+1}}{m+1} (1-\xi)^n \, d\xi$$

$$= \frac{n}{m+1} \int_0^1 \xi^{m+1} \left[ \frac{(1-\xi)^n}{n} \right] d\xi = \frac{n}{m+1} \int_0^1 \xi^{m+1} \left[ \int_0^{1-\xi} \eta^{n-1} d\eta \right] d\xi$$

$$= \frac{n}{m+1} \iint \xi^{m+1} \eta^{n-1} d\xi d\eta = \frac{n}{m+1} F(m+1, n-1)$$

Now we can set up the following sequence of formulas:

$$F(m,n) = \frac{n}{m+1} F(m+1,n-1) = \frac{n}{m+1} \frac{n-1}{m+2} F(m+2,n-2) = \ldots$$

$$= \frac{n(n-1) \ldots 2.1}{(m+1)(m+2) \ldots (m+n-1)(m+n)} F(m+n, 0)$$

$$= \frac{n(n-1) \ldots 2.1 \, . \, m(m-1) \, .. \, 2.1}{1.2. \ldots (m-1)m(m+1) \ldots (m+n)} F(m+n,0) = \frac{m! \, n!}{(m+n)!} F(m+n,0)$$

So only integrals of the form $F(m+n,0)$ are left to be calculated:

$$\iint \xi^{m+n} \, d\xi d\eta = \int_0^1 \xi^{m+n} \left[ \int_0^{1-\xi} d\eta \right] d\xi = \int_0^1 \xi^{m+n} (1-\xi) d\xi$$

$$= \int_0^1 \xi^{m+n} d\xi - \int_0^1 \xi^{m+n+1} d\xi = \left[ \frac{\xi^{m+n+1}}{m+n+1} \right]_0^1 - \left[ \frac{\xi^{m+n+2}}{m+n+2} \right]_0^1$$

$$= \frac{1}{m+n+1} - \frac{1}{m+n+2} = \frac{(m+n+2)-(m+n+1)}{(m+n+1)(m+n+2)}$$

$$= \frac{1}{(m+n+1)(m+n+2)} = F(m+n,0)$$

Hence:

$$F(m,n) = \frac{m!\,n!}{(m+n)!}\,\frac{1}{(m+n+1)(m+n+2)} = \frac{m!\,n!}{(m+n+2)!}$$

This is the final result:

$$\iint \xi^m \eta^n \, d\xi d\eta = \frac{m!\,n!}{(m+n+2)!}$$

Now let's calculate a few of these triangle moments.

$$\text{Area} = \iint dxdy = \iint d\xi d\eta \,\Delta = \frac{0!\,0!}{(0+0+2)!}\,\Delta = \Delta/2 = \frac{1}{2}(x_1 y_2 - x_2 y_1)$$

Since all (other) moments have to be divided by this area, the outcome of their integrals have to be multiplied with a factor $2/\Delta$ . A first order moment is:

$$\frac{\iint x\,dxdy}{\text{Area}} = 2/\Delta \iint (x_1\xi + x_2\eta)\,d\xi d\eta\,\Delta = 2x_1 \iint \xi\,d\xi d\eta + 2x_2 \iint \eta\,d\xi d\eta$$

$$= 2x_1\,\frac{1!\,0!}{(1+0+2)!} + 2x_2\,\frac{0!\,1!}{(0+1+2)!} = x_1\,2/6 + x_2\,2/6 = (x_1+x_2)/3$$

In very much the same way (replace $x$ by $y$) we can prove that:

$$\frac{\iint y\,dxdy}{\text{Area}} = (y_1+y_2)/3$$

Different though it seems, this is the same as the familiar result that the coordinates of the midpoint of a triangle equal one-third of the coordinates of the vertices:

$$\overline{x} = x_0 + 1/3\,[(x_1 - x_0) + (x_2 - x_0)] = (x_0 + x_1 + x_2)/3$$
$$\overline{y} = y_0 + 1/3\,[(y_1 - y_0) + (y_2 - y_0)] = (y_0 + y_1 + y_2)/3$$

Second order moments are:

$$2/\Delta \iint x^2\,dxdy \quad \text{and} \quad 2/\Delta \iint y^2\,dxdy \quad \text{and} \quad 2/\Delta \iint xy\,dxdy$$

It is sufficient to calculate only the last integral. Proper substitutions in $\overline{xy}$ will take care of the other two later on.

$$2/\Delta \iint xy\,dxdy = 2 \iint (x_1\xi + x_2\eta)(y_1\xi + y_2\eta)\,d\xi d\eta$$

$$= 2x_1y_1 \iint \xi^2\,d\xi d\eta + 2x_1y_2 \iint \xi\eta\,d\xi d\eta + 2x_2y_1 \iint \xi\eta\,d\xi d\eta + 2x_2y_2 \iint \eta^2\,d\xi d\eta$$

$$= 2x_1y_1\frac{2!\,0!}{(2+0+2)!}+2x_1y_2\frac{1!\,1!}{(1+1+2)!}+2x_2y_1\frac{1!\,1!}{(1+1+2)!}+2x_2y_2\frac{0!\,2!}{(0+2+2)!}$$

$$= 2x_1y_1.1/12 \;+\; 2x_1y_2.1/24 \;+\; 2x_2y_1.1/24 \;+\; 2x_2y_2.1/12$$

The result is:

$$\overline{xy} = (2x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2)/12$$

Substitute $x$ instead of $y$ herein:

$$\overline{xx} = (2x_1x_1 + x_1x_2 + x_2x_1 + 2x_2x_2)/12 \quad \Longrightarrow \quad \overline{x^2} = (x_1x_1 + x_1x_2 + x_2x_2)/6$$

Or the reverse: $y$ instead of $x$. Giving:

$$\overline{y^2} = (y_1y_1 + y_1y_2 + y_2y_2)/6$$

However, second order moments should be evaluated preferrably with respect to the midpoint:

$$\overline{xy} - \overline{x}\,\overline{y} = (2x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2)/12 - (x_1+x_2)/3.(y_1+y_2)/3$$

$$= (6x_1y_1 + 3x_1y_2 + 3x_2y_1 + 6x_2y_2 - 4x_1x_1 - 4x_1y_2 - 4x_2y_1 - 4x_2y_2)/36$$

$$= (2x_1x_1 - x_1y_2 - x_2y_1 + 2x_2y_2)/36$$

By proper substitution:

$$\overline{x^2} - \overline{x}^2 = (x_1x_1 - x_1y_2 + x_2y_2)/18 \quad \text{and} \quad \overline{y^2} - \overline{y}^2 = (y_1y_1 - y_1y_2 + y_2y_2)/18$$

The question remains to be answered why these triangle integrals are supposed to be so important. Well, virtually every domain in the plane can be thought as being built up from (small) triangles. This means that every integral over such a domain can effectively be thought as a (weighted) sum of integrals over nothing else but triangles ($k$):

$$\frac{\iint x^m y^n\,dxdy}{\iint dxdy} = \frac{\sum_k \left[\iint x^m y^n\,dxdy\right]_k}{\sum_k \left[\iint dxdy\right]_k} = \frac{\sum_k \left[\iint x^m y^n\,d\xi d\eta\;\Delta\right]_k}{\sum_k \left[\iint d\xi d\eta\;\Delta\right]_k}$$

$$= \sum_k \left[\iint x^m y^n\,d\xi d\eta\right]_k .\, w_k \qquad \text{where} \quad w_k = \frac{\Delta_k}{\frac{1}{2}\sum_i \Delta_i}$$

That is: the triangle moments are weighted with their individual areas, divided by the total area of the domain. This is the main reason why triangle moments are so useful: you can compose all other planar moments out of them (within certain accuracy bounds). Provided that you are not too punctilious, anything else will not be needed.

## Boundary Moments

Hithereto, moments actually have been calculated over the surface which is enclosed by a contour. It's also possible, however, to define the moments of a contour over its boundary only. Since all contours can be thought as being built up from infinitely many straight line segments, it is sufficient, for our purpose, to restrict attention to straight line segments. The governing equations of any such a segment are:

$$x = x_1 + (x_2 - x_1).t$$
$$y = y_1 + (y_2 - y_1).t$$

Here $(x_1, y_1)$ and $(x_2, y_2)$ are the end-points of the segment and $0 \le t \le 1$ .
The moments are calculated now in chronological order. First we do the *length* of the line segment:

$$\int_{(1)}^{(2)} ds = \int_{(1)}^{(2)} \sqrt{dx^2 + dy^2} = \int_0^1 \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\, dt \quad \Longrightarrow$$

$$\text{Length} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Then we do the *middle* of the line segment:

$$\frac{\left(\int_{(1)}^{(2)} x\, ds\ ,\ \int_{(1)}^{(2)} y\, ds\right)}{\text{Length}} = \frac{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\left(\int_0^1 x\, dt\ ,\ \int_0^1 y\, dt\right)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}$$

$$= \left(\int_0^1 x\, dt\ ,\ \int_0^1 y\, dt\right)$$

It is seen that the square root, representing the length of the line segment, is just cancelled out. This is a common feature of all boundary moment calculations.

$$\left(\int_0^1 x\, dt\ ,\ \int_0^1 y\, dt\right) = \left(\int_0^1 [x_1 + (x_2 - x_1).t]\ dt\ ,\ \int_0^1 [y_1 + (y_2 - y_1).t]\ dt\right)$$

$$= \left(x_1.1 + (x_2 - x_1).\frac{1}{2}\ ,\ y_1.1 + (y_2 - y_1).\frac{1}{2}\right) \quad \Longrightarrow$$

$$\text{Middle} = \frac{1}{2}\,(x_1 + x_2\ , y_1 + y_2)$$

Well, guess that we did'nt expect this result. Anyway, here comes the matrix of second order moments, at last:

$$\begin{pmatrix} \int_0^1 x^2\, dt & \int_0^1 xy\, dt \\ \int_0^1 xy\, dt & \int_0^1 y^2\, dt \end{pmatrix}$$

Same trick as always: calculate $\overline{xy}$ only and find $\overline{x^2}$ as well as $\overline{y^2}$ later on by proper substitution herein.

$$\int_0^1 xy\,dt = \int_0^1 \left[x_1 + (x_2 - x_1).t\right]\left[y_1 + (y_2 - y_1).t\right]\,dt$$

$$= x_1 y_1 + x_1(y_2 - y_1).\frac{1}{2} + (x_2 - x_1)y_1.\frac{1}{2} + (x_2 - x_1)(y_2 - y_1).\frac{1}{3}$$

Here come the results:

$$\overline{xy} = \frac{1}{3}(x_1 y_1 + x_2 y_2) + \frac{1}{6}(x_1 y_2 + x_2 y_1) \quad \Longrightarrow$$

$$\overline{x^2} = \frac{1}{3}(x_1 x_1 + x_1 x_2 + x_2 x_2) \quad \text{and} \quad \overline{y^2} = \frac{1}{3}(y_1 y_1 + y_1 y_2 + y_2 y_2)$$

Boundary moments (**lijn**) are used in the program as *error estimates* for the corresponding surface moments (**vlak**). To see how this is possible, some well-known facts about the area and (arc)length of a circle come into mind:

$$\text{Area} = \pi r^2 \qquad \text{Length} = 2\pi r \quad \Longrightarrow \quad d\text{Area} = 2\pi r\,dr = \text{Length.Thickness}$$

The same kind of thought can be produced for a square with edge-length $2a$ :

$$\text{Area} = (2a)^2 \qquad \text{Length} = 8a \quad \Longrightarrow \quad d\text{Area} = 8a\,da = \text{Length.Thickness}$$

The generalization of these facts reads as follows:

$$d(\text{surface moment}) = (\text{boundary moment}) \times \text{thickness}$$

In words: an estimate for the error in a surface moment is the magnitude of the corresponding boundary moment times the thickness of the boudary. The latter is not difficult to find, since most of the time the accuracy / thickness of the boundary is equal to one or two pixels in the image at hand.

## Asymmetric 2-D Gauss function

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, $\sigma_{xx}$ 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 $\lambda_1$ and $\lambda_2$. 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:

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

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

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:

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

## Ellipses of Inertia

A measure for the width of the bell-shaped Gauss curve in one dimension is the *spread* called $\sigma$ :

$$\sigma = \sqrt{\sigma_{xx}} \qquad \text{in} \quad g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x^2/\sigma^2)} \quad \Longrightarrow \quad g(\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}}$$

Analogously, a sensible "width" of the 2-D skewed Gaussian distribution could be conceived as the value of the exponent which is such that:

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

$$= \frac{1}{2\pi\sqrt{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2}}\, e^{-\frac{1}{2}} \quad \Longrightarrow$$

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

Thus, instead of being characterized by a single number, the "size" of the 2-D skewed Gaussian distribution is bound by a *curve* of the form $F(x,y) = 1$. This outcome is far more transparent in eigenvector coordinates, where the same function is written as:

$$(x-\mu_x)^2/\lambda_1 + (y-\mu_y)^2/\lambda_2 = 1$$

Substitute:

$$\begin{array}{ll} x := x - \mu_x & a^2 := \lambda_1 \\ x := y - \mu_y & b^2 := \lambda_2 \end{array}$$

Then the equation for the curve $F(x,y) = 1$ is:

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

This means that the area of dominant values for the skewed 2-D Gauss function is bounded by an ellipse with (its midpoint in the centre) and its (half) axes equal to the square roots of the main moments of inertia.

$$a := \sqrt{\lambda_1} \quad \text{and} \quad b := \sqrt{\lambda_2}$$

Reason why this ellipse is commonly called the *ellipse of inertia*. Redefine the function $F(x,y)$ as:

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

Some general formulas for the (partial) differentiation of implicit functions are:

$$0 = dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy \quad \Longrightarrow \quad \frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} \quad \text{and} \quad \frac{dx}{dy} = -\frac{\partial F/\partial y}{\partial F/\partial x}$$

Enabling us to find tangent lines, as follows:

$$\frac{dy}{dx} = 0 \quad \Longleftrightarrow \quad \frac{\partial F}{\partial x} = 0 \qquad \text{and} \qquad \frac{dx}{dy} = 0 \quad \Longleftrightarrow \quad \frac{\partial F}{\partial y} = 0$$

Applied to the curve at hand:

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

First do $dy/dx$ for finding the horizontal tangents:

$$\frac{\partial F}{\partial x} = 0 \quad \Longrightarrow \quad \sigma_{yy}(x-\mu_x) - \sigma_{xy}(y-\mu_y) = 0 \quad \Longrightarrow \quad (x-\mu_x) = \frac{\sigma_{xy}}{\sigma_{yy}}(y-\mu_y)$$

Substitute into $F(x,y) = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$ :

$$\sigma_{yy}\left[\frac{\sigma_{xy}}{\sigma_{yy}}(y-\mu_y)\right]^2 - 2\sigma_{xy}\left[\frac{\sigma_{xy}}{\sigma_{yy}}(y-\mu_y)\right](y-\mu_y) + \sigma_{xx}(y-\mu_y)^2 = \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$$

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

$$\Longleftrightarrow \quad \frac{(y-\mu_y)^2}{\sigma_{yy}} = 1 \quad \Longleftrightarrow \quad y = \mu_y \pm \sqrt{\sigma_{yy}}$$

$$\text{and} \quad (x-\mu_x) = \frac{\sigma_{xy}}{\sigma_{yy}}(\pm\sqrt{\sigma_{yy}}) \quad \Longrightarrow \quad x = \mu_x \pm \frac{\sigma_{xy}}{\sqrt{\sigma_{yy}}}$$

In very much the same way we find the vertical tangents:

$$\frac{dx}{dy} = 0 \quad \Longleftrightarrow \quad x = \mu_x \pm \sqrt{\sigma_{xx}} \quad \text{and} \quad y = \mu_y \pm \frac{\sigma_{xy}}{\sqrt{\sigma_{xx}}}$$

When visualizing the result, it is seen that the ellipse of inertia is enclosed by a *Bounding Box* which is defined by the diagonal:

$$(\mu_x - \sqrt{\sigma_{xx}}, \mu_y - \sqrt{\sigma_{yy}}) - (\mu_x + \sqrt{\sigma_{xx}}, \mu_y + \sqrt{\sigma_{yy}})$$

Because Gauss functions are expensive to compute, it is desirable to have an estimate of the cut-off value, beyond which the values of Gauss function can be safely set to zero. Such is the case if, beyond certain values for $(x,y)$:

$$e^{-\frac{1}{2}\left[\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2\right]/(\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2)} < \epsilon \quad \Longleftrightarrow$$

$$-\frac{1}{2}\frac{\sigma_{yy}(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_{xx}(y-\mu_y)^2}{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2} < ln(\epsilon) \quad \Longleftrightarrow$$

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

Again, the outcome is far more transparent in eigenvector coordinates, where:

$$(x-\mu_x)^2/\lambda_1 + (y-\mu_y)^2/\lambda_2 > 2\,ln(1/\epsilon)$$

Substitute:

$$\begin{array}{ll} x := x - \mu_x & a^2 := \lambda_1.2\,ln(1/\epsilon) \\ x := y - \mu_y & b^2 := \lambda_2.2\,ln(1/\epsilon) \end{array}$$

Then the above condition for the uninteresting area can be written as:

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 > 1$$

This means that the area of interesting values for the Gauss function is bounded by still another ellipse of inertia with (half) axes:

$$a := \sqrt{\lambda_1.2\,ln(1/\epsilon)}$$

$$b := \sqrt{\lambda_2.2\,ln(1/\epsilon)}$$

Hence it is possible to define a (Pascal) function for the part of the plane where values of the skewed 2-D Gauss function are worthwile to be calculated, under the assumption that the vector *mu* and the tensor *sigma* are known globally:

```
function interesting(x,y : double) : boolean;
var
  Det : double;
begin
  Det := sigma.xx*sigma.yy - sqr(sigma.xy);
  interesting :=
  ( sigma.yy*sqr(x-mu.x) - 2*sigma.xy*(x-mu.x)*(y-mu.x)
  + sigma.xx*sqr(y-mu.y) < Det * 2*ln(1/epsilon) );
end;
```

Sometimes, it may be necessary to have kind of an estimate of the area which is covered by an ellipse of inertia, for the reason that the number of points (i.e.pixels) inside the area of interest is proportional to such an area. One could, for example, carry out a *Flood Fill* on "black" pixels inside. It is known that the area of an ellipse with axes $a$ and $b$ is given by:

$$\pi\, a.b = \pi\,\sqrt{\lambda_1}\sqrt{\lambda_2} = \pi\,\sqrt{\lambda_1\lambda_2} = ln(1/\epsilon)\; 2\pi\,\sqrt{\sigma_{xx}\sigma_{yy}-\sigma_{xy}^2}$$

Hence the area of interest is exactly equal to the norming factor, times the natural logarithm of one divided by the desired accuracy. The Bounding Box diagonal of the ellipse of interest is given by:

$$(\mu_x - \sigma_x, \mu_y - \sigma_y) - (\mu_x + \sigma_x, \mu_y + \sigma_y)$$

$$\text{where} \quad \sigma_x := \sqrt{\sigma_{xx}.2\,ln(1/\epsilon)} \quad \text{and} \quad \sigma_y := \sqrt{\sigma_{yy}.2\,ln(1/\epsilon)}$$

Resulting in a BB area which exceeds the area of the ellipse with a factor greater than $> 8/2\pi$ .

## IBM PunchCard

Non ASCII: a
Non Fortran: f
Card Punch: O

```
    f         a f            f   f         a ffff           ff  f
    &ABCDEFGHI .<(+-JKLMNOPQR!$*);0/STUVWXYZ ,%_>^ 123456789:#@'="
   -----------------------------------------------------------------
2 | OOOOOOOOOOOOOOO                                                 | 2
1 |                OOOOOOOOOOOOOOO                                  | 1
0 |                               OOOOOOOOOOOOOOOO                  | 0
1 |  O              O              O               O                | 1
2 |   O       O      O       O      O       O       O       O       | 2
3 |    O       O      O       O      O       O       O       O      | 3
4 |     O       O      O       O      O       O       O       O     | 4
5 |      O       O      O       O      O       O       O       O    | 5
6 |       O       O      O       O      O       O       O       O   | 6
7 |        O              O              O       O       O       O  | 7
8 |         O OOOOO        O OOOOO        O OOOOOO        O OOOOOO  | 8
9 |          O              O              O               O        | 9
   -----------------------------------------------------------------
    &ABCDEFGHI .<(+-JKLMNOPQR!$*);0/STUVWXYZ ,%_>^ 123456789:#@'=?
```

Sample content:

```
      SUBROUTINE GELB(R,A,M,N,MUD,MLD,EPS,IER)
      IF(MAX.GT.NN) STOP "DUAAL: TOO MANY NODES"
   10 IF(MAX.GT.NN) CALL DISPLA("ELEMENT =",NE)
      PROGRAM AFRIKA(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
   30 WRITE(6,3) LAND,(DUAL(LAND,PUNT),PUNT=1,10)
      COMMON /TOPOL/ NDS,NO(1)  $  COMMON /DUAAL/ NN,DUAL(1)
```