Lie Groups and Variances

Author: Han de Bruijn
Date: 2006 June

Continuous linear transformations in one and two dimensions are considered. It is argued that they all can be thought as being composed from six basic ones: a translation and/or scaling in the x and y direction, a rotation and a skewing. The first and second order moments of a points cloud are defined. And next it is studied how these transform with the above. Most interesting is the complete analogy between rotation and skewing. Four quantities play a predominant role: $\sigma_{xx} + \sigma_{yy}$, $\sigma_{xx} - \sigma_{yy}$, $2\sigma_{xy}$ and $\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2$. The second and the third act like a vector while the first is invariant, for rotation. The first and the third act like a vector while the second is invariant, for skewing. The last quantity is invariant for both rotation and skewing.

Lie Groups 1-D

Any differentiable function can be developed into a Taylor series around one of its points:

$$f(x+a) = f(a) + x.f'(a) + \frac{1}{2}x^2.f''(a) + ...$$

Here the choice of the point a is arbitrary. A Taylor series expansion around point zero is quite common:

$$f(x) = f(0) + x.f'(0) + \frac{1}{2}x^2.f''(0) + ...$$

But, the other way around, one could also write:

$$f(x+a) = f(x) + a.f'(x) + \frac{1}{2}a^2.f''(x) + ...$$

With physical systems, there is no preferred origin. This means that, always, a function f_a can be defined, which is merely shifted with respect to f with a certain amount a :

$$f_a(x) = f(x+a) = f(x) + a.f'(x) + \frac{1}{2}a^2.f''(x) + \frac{1}{3!}a^3 f'''(x) + ...$$

We thus see that the shifted *function* may be expanded as a Taylor series around the original function. It is emphasized that developing a function as a power series around another function is *quite different* from developing a function as a power series around one of its own values, though the formulas are more or less similar. The series expansion of the shifted function can also be written as follows: $f_a(x) = f(x+a) =$

$$\left[1 + a.\frac{d}{dx} + \frac{1}{2}\left(a.\frac{d}{dx}\right)^2 + \frac{1}{3!}\left(a.\frac{d}{dx}\right)^3 + ... + \frac{1}{n!}\left(a.\frac{d}{dx}\right)^n + ...\right] f(x)$$

$$\Longrightarrow \quad f_a(x) = f(x+a) = e^{a.\frac{d}{dx}} f(x)$$

Thus the shifted function is obtained by taking into account the effect of the *translation operator* $\exp(a.d/dx)$ as it acts upon $f(x)$.
Instead of translating a function over a certain distance, which seems to be a rather trivial operation anyway, let us consider *scaling*. This means that we are going to make intervals of the independent variable smaller, or larger, with a factor $\lambda > 0$. The transformed function is then defined by:

$$f_\lambda(x) = f(\lambda.x)$$

Like with translations, it would be nice to develop the function $f_\lambda(x)$ into a Taylor series expansion around the original $f(x)$. But this is not as simple as in the former case. Unless a clever trick is devised, which reads as follows. Define a couple of new variables, a and y, and a new function g:

$$\lambda = e^a \quad \text{and} \quad x = e^y \quad \text{and} \quad g(y) = f(e^y)$$

Then, indeed, we can develop something into a Taylor series:

$$f_\lambda(x) = f(e^a.e^y) = f(e^{a+y}) = g(y+a) = e^{a.\frac{d}{dy}} g(y)$$

A variable such as y, which renders the transformation to be resemblant to a translation, is commonly called a *canonical* variable. In the case of a scaling transformation, the canonical variable is obtained by taking the logarithm of the independent variable: $y = ln(x)$. Working back to the original variables and the original function:

$$g(y) = f(e^y) = f(x) \quad \text{and} \quad a = ln(\lambda)$$

$$\frac{d}{dy} = \frac{dx}{dy}\frac{d}{dx} = e^y \frac{d}{dx} = x\frac{d}{dx}$$

Where the operator $x.d/dx$ is commonly called the *infinitesimal operator* of a scaling transformation. An infinitesimal operator always equals differentiation to the canonical variable, which converts the transformation into a translation. We have already met, of course, the infinitesimal operator for the translations themselves, which is simply given by (d/dx). This all leads, rather quickly, to the following somewhat less-trivial result:

$$f_\lambda(x) = f(\lambda.x) = e^{ln(\lambda).x\frac{d}{dx}} f(x)$$

Written out as a "true" Taylor series:

$$f(\lambda.x) = f(x) + ln(\lambda).x\frac{df}{dx} + \frac{1}{2}ln^2(\lambda).x\frac{d(x.df/dx)}{dx} + ...$$

However, there is something bogus about this argument, since it is obvious that the result has only be validated here for $x > 0$. In order to be certain that things are also valid for negative values of x , we must actually carry out the calculation. It is sufficient to do this for the scaling transformation of x itself, which is represented by the series $exp(ln(\lambda).x.d/dx)x$:

$$e^{ln(\lambda).x\frac{d}{dx}}x = x + ln(\lambda).x\frac{dx}{dx} + \frac{1}{2}ln^2(\lambda).x\frac{d(x.dx/dx)}{dx} + ...$$

$$= \left[1 + ln(\lambda) + \frac{1}{2}ln^2(\lambda) + ...\right] x = e^{ln(\lambda)}x = \lambda x$$

Herewith the series expansion is verified for all real values of x . However, since λ must be positive, there exists no continuous transition towards problems where values are, at the same time, inverted or *mirrorred*, like in:

$$f_\lambda(x) = f(-\lambda.x)$$

For this to happen, the scaling transformation would to have to pass through a point where everything is contracted to zero:

$$f_\lambda(x) = f(0.x)$$

This already reveals a glimpse of the *topological issues* which may be associated with the notion of Taylor series around a function. These have been troubling the area of research ever since. (Quite unneccessarily in my opinion, but that's another matter.)

Lie Groups 2-D

An example of a Continuous transformation in two dimensions is Rotation over an angle θ:

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

It might be asked how the rotation of the independent variables works out for a function of these variables. With other words, how the following function would be expanded as a Taylor series expansion around the original $f(x, y)$:

$$f_\theta(x, y) = f(x_\theta, y_\theta) = f(\cos(\theta)x - \sin(\theta)y\,,\, \sin(\theta)x + \cos(\theta)y)$$

Define other (polar) variables (r, ϕ) as:

$$x = r.\cos(\phi) \quad \text{and} \quad y = r.\sin(\phi)$$

Giving for the transformed variables:

$$x_\theta = r.\cos(\phi).\cos(\theta) - r.\sin(\phi).\sin(\theta) = r.\cos(\phi + \theta)$$
$$y_\theta = r.\cos(\phi).\sin(\theta) + r.\sin(\phi).\cos(\theta) = r.\sin(\phi + \theta)$$

We see that ϕ is a proper canonical variable. Another function $g(\psi)$ is defined with this canonical variable as the independent one:

$$g(\phi) = f(r.\cos(\phi), r.\sin(\phi)) = f(x, y)$$

Now rotating $f(x,y)$ over an angle θ corresponds with a translation of $g(\phi)$ over a distance θ. Therefore $g(\phi+\theta)$ can be developed into a Taylor series around the original:

$$g(\phi+\theta) = g(\phi) + \theta.\frac{dg(\phi)}{d\phi} + \frac{1}{2}\theta^2.\frac{d^2g}{d\phi^2} + ...$$

Working back to the original variables (x,y) with a well known chain rule for partial derivatives:

$$\frac{dg}{d\phi} = \frac{\partial g}{\partial x}\frac{dx}{d\phi} + \frac{\partial g}{\partial y}\frac{dy}{d\phi}$$

Where:

$$\frac{dx}{d\phi} = -r.\sin(\phi) = -y \quad \text{and} \quad \frac{dy}{d\phi} = +r.\cos(\phi) = +x \quad \Longrightarrow$$

$$\frac{dg}{d\phi} = \frac{\partial g}{\partial x}.(-y) + \frac{\partial g}{\partial y}.(+x) \quad \Longrightarrow \quad \frac{d}{d\phi} = x.\frac{\partial}{\partial y} - y.\frac{\partial}{\partial x}$$

Herewith we find that the operator $(x.\frac{\partial}{\partial y} - y.\frac{\partial}{\partial x})$ is the *infinitesimal operator* for Plane Rotations. It is equal to differentiation with respect to the canonical variable, as expected. The end-result is:

$$f_\theta(x,y) = \sum_{k=0}^{\infty}\frac{1}{k!}\left[\theta(x.\frac{\partial}{\partial y} - y.\frac{\partial}{\partial x})\right]^k f(x,y) = e^{\theta(x\partial/\partial y - y\partial/\partial x)}f(x,y)$$

This is true for *any* function $f(x,y)$. In particular, the independent variables themselves can be conceived as such functions. Which means that:

$$x_\theta = e^{\theta(x\partial/\partial y - y\partial/\partial x)}x \quad \text{and} \quad y_\theta = e^{\theta(x\partial/\partial y - y\partial/\partial x)}y$$

It is easily demonstrated that:

$$(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})x = -y \quad \text{and} \quad (x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})y = x$$

Herewith we find:

$$\sum_{k=0}^{\infty}\left[\theta(x.\frac{\partial}{\partial y} - y.\frac{\partial}{\partial x})\right]^k x = 1 - \theta.y - \frac{1}{2}\theta^2.x + \frac{1}{3!}\theta^3.y + \frac{1}{4!}\theta^4.x + ...$$

$$= \cos(\theta)x - \sin(\theta)y = x_\theta$$

Likewise we find:

$$\sum_{k=0}^{\infty}\left[\theta(x.\frac{\partial}{\partial y} - y.\frac{\partial}{\partial x})\right]^k y = 1 + \theta.x - \frac{1}{2}\theta^2.y - \frac{1}{3!}\theta^3.x + \frac{1}{4!}\theta^4.y + ...$$

$$= \sin(\theta)x + \cos(\theta)y = y_\theta$$

Thus, indeed, the formulas for a far-form-infinitesimal rotation over an finite angle θ can be reconstructed from the expansions.
Now suppose that we change a minor thing in the infinitesimal operator for plane rotations, namely the minus sign into a plus sign, giving:

$$\frac{d}{dp} = x.\frac{\partial}{\partial y} + y.\frac{\partial}{\partial x}$$

And it is wondered what global transformation will come out of such a subtle change. Calculate in very much the same way:

$$x_p = \sum_{k=0}^{\infty} \left[p(x.\frac{\partial}{\partial y} + y.\frac{\partial}{\partial x}) \right]^k x \quad \text{and} \quad y_p = \sum_{k=0}^{\infty} \left[p(x.\frac{\partial}{\partial y} + y.\frac{\partial}{\partial x}) \right]^k y$$

It is easily demonstrated that:

$$(x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x})x = y \quad \text{and} \quad (x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x})y = x$$

Herewith we find:

$$\sum_{k=0}^{\infty} \left[p(x.\frac{\partial}{\partial y} + y.\frac{\partial}{\partial x}) \right]^k x = 1 + p.y + \frac{1}{2}p^2.x + \frac{1}{3!}p^3.y + \frac{1}{4!}p^4.x + ...$$

$$= \cosh(p)x + \sinh(p)y = x_p$$

Likewise we find:

$$\sum_{k=0}^{\infty} \left[p(x.\frac{\partial}{\partial y} + y.\frac{\partial}{\partial x}) \right]^k y = 1 + p.x + \frac{1}{2}p^2.y + \frac{1}{3!}p^3.x + \frac{1}{4!}p^4.y + ...$$

$$= \sinh(p)x + \cosh(p)y = y_p$$

For those who don't know what the meaning is of a hyperbolic cosine $\cosh(p)$ and a hyperbolic sine $\sinh(p)$, the definitions are:

$$\cosh(p) = \frac{e^{+p} + e^{-p}}{2} \quad \text{and} \quad \sinh(p) = \frac{e^{+p} - e^{-p}}{2}$$

The truth of the above can easily be checked out now. What we have found here is the following transformation:

$$\begin{cases} x_p = \cosh(p)x + \sinh(p)y \\ y_p = \sinh(p)x + \cosh(p)y \end{cases}$$

In structural mechanics, this beasty is known as the *Deformation Tensor* and, as such, it should have been known for a some time. But it really became a famous transformation when the theory of Special Relativity was discovered, where it is known as the Lorentz Transformation. A special property of pure deformations, like with rotations, is that the determinant of the matrix is always equal to unity:

$$\left\| \begin{array}{cc} \cosh(p) & \sinh(p) \\ \sinh(p) & \cosh(p) \end{array} \right\| = \cosh^2(p) - \sinh^2(p) = 1$$

Meaning that pure deformations leave volumes of the material unchanged.
This more or less completes our discussion about transformations in one and two dimensions. We have found six infinitesimal operators in total, if we also take into account both one-dimensional (x, y) directions in the plane. Here comes a list of differentiations to the canonical variables, also called infinitesimal transformations, which have been found so far.

$$\begin{array}{rc} \text{Translation x:} & \partial/\partial x \\ \text{Translation y:} & \partial/\partial y \\ \text{Scaling for x:} & \partial/\partial ln(x) = x.\partial/\partial x \\ \text{Scaling for y:} & \partial/\partial ln(y) = y.\partial/\partial y \\ \text{Flat rotation:} & d/d\theta = x.\partial/\partial y - y.\partial/\partial x \\ \text{Flat skewness:} & d/dp = x.\partial/\partial y + y.\partial/\partial x \end{array}$$

What we are talking about here all the time is the so-called *Theory of Lie Groups*. Most of the time, Lie Groups have been the kind of mathematics which tend to be incomprehensible for most human beings. Maybe that's because it is embedded, nowadays, a great deal, in Physics on Elementary Particles and other SF-like subjects. It has been shown here that very basic, 1- and 2-dimensional examples provide useful material to demonstrate the whole idea, *the gist* of the Theory, while the more difficult, "advanced" topics can be left aside.

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:

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

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

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

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

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 μ_x or all (squares of the) spreads with respect to the mean. Then we can drop the dependence on (p) in:

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

Translations can be handled in one dimension. Let $x_i := x_i + a$ in:

$$\overline{x}\,' = \sum_k w_k (x_k + a) = \sum_k w_k x_k + a \sum_k w_k = \overline{x} + a$$

Second order moment (with respect to $\sum_k w_k x_k = 0$):

$$\sigma'_{xx} = \sum_k w_k (x_k + a)^2 = \sum_k w_k x_k^2 + a \sum_k x_k + a^2 = \sigma_{xx} + a^2$$

Which is the old variance plus the variance of the midpoint.
Scaling can be handled in one dimension as well. Let $x_i := \lambda x_i$ in:

$$\sigma'_{xx} = \sum_k w_k (\lambda x_k)^2 = \lambda^2 \sigma_{xx} \quad \Longrightarrow \quad \sqrt{\sigma'_{xx}} = \lambda \sqrt{\sigma_{xx}}$$

As expected, the spread scales in the same way as the coordinates.

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 σ_{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 σ_{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 σ_{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 (μ_x, μ_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 σ_{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 ρ is the so-called *cross-correlation coefficient*.
With respect to translation over (a, b) and scaling with (λ, μ), the midpoints and variances transform as follows:

$$\overline{x}\,' = \overline{x} + a \qquad ; \qquad \overline{y}\,' = \overline{y} + b$$

$$\overline{x^2}\,' = \overline{x^2} + a^2 \qquad ; \qquad \overline{y^2}\,' = \overline{y^2} + b^2 \qquad ; \qquad \overline{xy}\,' = \overline{xy} + ab$$

$$\overline{x^2}\,' = \lambda^2 \overline{x^2} \qquad ; \qquad \overline{y^2}\,' = \mu^2 \overline{y^2} \qquad ; \qquad \overline{xy}\,' = \lambda\mu\overline{xy}$$

Rotations and Skewing

Let's see how the second order moments behave under rotations and skewing. Start with:

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

Then:

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

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

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

And:

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

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

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

Last but not least:

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

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

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

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

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

Two trigonometric formulas should me memorized here:

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

Herewith:

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

And:

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

Addition simply results in:

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

Which means that the *trace* (: sum of the diagonal elements) of the tensor of inertia is invariant for rotations. Substraction gives something more subtle, but likely interesting:

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

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

On the other hand we have:

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

See the structure? In matrix form:

$$\left[\begin{array}{c} \sigma'_{xx} - \sigma'_{yy} \\ 2\sigma'_{xy} \end{array}\right] \left[\begin{array}{cc} \cos(2\theta) & \sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{array}\right] \left[\begin{array}{c} \sigma_{xx} - \sigma_{yy} \\ 2\sigma_{xy} \end{array}\right]$$

With other words: the quantity $[\sigma_{xx} - \sigma_{yy}, 2\sigma_{xy}]$ transforms like a vector, but with *double* the rotation angle. As a consequence, if this vector is zero, then there is *no way* to make it nonzero with help of a rotation. This situation occurs more often than one might think: namely with figures that are *symmetric*, for example a square, or a circle, or even the letter 'A'.
Another consequence is:

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

$$(\sigma_{xx} + \sigma_{yy})^2 - 4\sigma_{xx}\sigma_{xx} + (2\sigma_{xy})^2 = \text{invariant} \quad \Longrightarrow$$

$$\sigma_{xx}\sigma_{xx} - \sigma_{xy}^2 = \text{invariant}$$

Since it is already known that the trace $\sigma_{xx} + \sigma_{yy}$ is invariant for rotations. It is seen that $\sigma_{xx}\sigma_{xx} - \sigma_{xy}^2$ is the *determinant* of the inertial tensor:

$$\left[\begin{array}{cc} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{array}\right]$$

So far so good for rotations. How about skewing? Well, this turns out to be entirely analogous. So we did a copy / paste and changed circle functions into hyperbolic functions everywhere. Start with:

$$\left\{ \begin{array}{l} x' = \cosh(p)x + \sinh(p)y \\ y' = \sinh(p)x + \cosh(p)y \end{array} \right.$$

Then:

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

$$= \cosh^2(p)\sum_k w_k x_k^2 + 2\sinh(p)\cosh(p)\sum_k w_k x_k y_k + \sinh^2(p)\sum_k w_k y_k^2$$

$$\Longrightarrow \quad \sigma'_{xx} = \cosh^2(p)\sigma_{xx} + 2\sinh(p)\cosh(p)\sigma_{xy} + \sinh^2(p)\sigma_{yy}$$

And:

$$\sigma'_{yy} = \sum_k w_k (y'_k)^2 = \sum_k w_k \left[\sinh(p) x_k + \cosh(p) y_k\right]^2$$

$$= \sinh^2(p) \sum_k w_k x_k^2 + 2\sinh(p)\cosh(p) \sum_k w_k x_k y_k + \cosh^2(p) \sum_k w_k y_k^2$$

$$\Longrightarrow \quad \sigma'_{yy} = \sinh^2(p)\sigma_{xx} + 2\sinh(p)\cosh(p)\sigma_{xy} + \cosh^2(p)\sigma_{yy}$$

Last but not least:

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

$$\sum_k w_k \left[\cosh(p) x_k + \sinh(p) y_k\right]\left[\sinh(p) x_k + \cosh(p) y_k\right] =$$

$$\cosh(p)\sinh(p) \sum_k w_k x_k^2 + \sinh(p)\cosh(p) \sum_k w_k y_k^2$$

$$+ \left[\cosh^2(p) + \sinh^2(p)\right] \sum_k w_k x_k y_k$$

$$\Longrightarrow \quad \sigma'_{xy} \cosh(p)\sinh(p)\left(\sigma_{yy} + \sigma_{xx}\right) + \left[\cosh^2(p) + \sinh^2(p)\right]\sigma_{xy}$$

Two hyperbolic function formulas should me memorized here:

$$\sinh(2p) = 2\sinh(p)\cosh(p) \quad \text{and} \quad \cosh(2p) = \cosh^2(p) + \sinh^2(p)$$

Herewith:

$$\sigma'_{xy} = \frac{1}{2}\sinh(2p)(\sigma_{xx} + \sigma_{yy}) + \cosh(2p)\sigma_{xy}$$

And:

$$\sigma'_{xx} = \cosh^2(p)\sigma_{xx} + \sinh(2p)\sigma_{xy} + \sinh^2(p)\sigma_{yy}$$
$$\sigma'_{yy} = \sinh^2(p)\sigma_{xx} + \sinh(2p)\sigma_{xy} + \cosh^2(p)\sigma_{yy}$$

With $\cosh^2(p) - \sinh^2(p) = 1$, substraction simply results in:

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

Addition gives something more subtle, but likely interesting: $\sigma'_{xx} + \sigma'_{yy} =$

$$\left[\cosh^2(p) + \sinh^2(p)\right]\sigma_{xx} + 2\sinh(2p)\sigma_{xy} + \left[\sinh^2(p) + \cosh^2(p)\right]\sigma_{yy}$$

$$\Longrightarrow \quad \sigma'_{xx} + \sigma'_{yy} = \cosh(2p)(\sigma_{xx} + \sigma_{yy}) + \sinh(2p)2\sigma_{xy}$$

On the other hand we have:

$$2\sigma'_{xy} = \sinh(2p)(\sigma_{xx} + \sigma_{yy}) + \cosh(2p)2\sigma_{xy}$$

See the structure? In matrix form:

$$\begin{bmatrix} \sigma'_{xx} + \sigma'_{yy} \\ 2\sigma'_{xy} \end{bmatrix} \begin{bmatrix} \cosh(2p) & \sinh(2p) \\ -\sinh(2p) & \cosh(2p) \end{bmatrix} \begin{bmatrix} \sigma_{xx} + \sigma_{yy} \\ 2\sigma_{xy} \end{bmatrix}$$

With other words: the quantity $[\sigma_{xx} + \sigma_{yy}, 2\sigma_{xy}]$ transforms like a vector, but with *double* the hyperbolic angle. It will never happen that this vector is zero, though, because $\sigma_{xx} + \sigma_{yy} = 0$ would mean that the points cloud considered has no size at all. Whatever. Another consequence is:

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

$$(\sigma_{xx} - \sigma_{yy})^2 + 4\sigma_{xx}\sigma_{xx} - (2\sigma_{xy})^2 = \text{invariant} \quad \Longrightarrow$$

$$\sigma_{xx}\sigma_{xx} - \sigma_{xy}^2 = \text{invariant}$$

Since it is already known that the quantity $\sigma_{xx} - \sigma_{yy}$ is invariant for skewing. It is seen that $\sigma_{xx}\sigma_{xx} - \sigma_{xy}^2$ is the same *determinant* of the inertial tensor, which thus is invariant for *both* rotations and skewing.

Disclaimers

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