UNIVERSITY OF MINNESOTA

Statistics 5401 September 26, 2005

The Shapes of Ellipses and Ellipsoids

An ellipse centered at (w_{10}, w_{20}) in the plane with coordinates (w_1, w_2) consists of all points (w_1, w_2) such that

$$q^{11}(w_1-w_{10})^2 + 2q^{12}(w_1-w_{10})(w_2-w_{20}) + q^{22}(w_2-w_{20})^2 = K^2.$$

Here q^{11}, q^{12}, and q^{22} are constants that determine the *shape* of the ellipse, and K^2 is a constant that determines the *size* of the ellipse.

The entire ellipse, including its interior, is the set of points such that

$$q^{11}(w_1-w_{10})^2 + 2q^{12}(w_1-w_{10})(w_2-w_{20}) + q^{22}(w_2-w_{20})^2 \leq K^2$$

For these equations to define an ellipse, the q^{ij} must satisfy

$$q^{11} > 0,\ q^{22} > 0 \text{ and } q^{11}q^{22} > (q^{12})^2.$$

The inequalities $q^{11} > 0$, $q^{22} > 0$ and $q^{11}q^{22} > (q^{12})^2$ are equivalent to the condition that the matrix

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} = \begin{bmatrix} q^{11} & q^{12} \\ q^{12} & q^{22} \end{bmatrix}^{-1}$$

and its inverse $\mathbf{Q}^{-1}$ are positive definite.

You can generalize the concept of an ellipse to $p > 2$ dimensions. Suppose $\mathbf{Q} = \mathbf{Q}' = [q_{ij}]$ is a p by p *positive definite symmetric* matrix and $\mathbf{w}_0 = [w_{10},\ldots,w_{p0}]'$ is a p-vector. Then the set of p-dimensional vectors

$$\{\mathbf{w} \mid (\mathbf{w} - \mathbf{w}_0)'\mathbf{Q}^{-1}(\mathbf{w} - \mathbf{w}_0) = K^2\}$$

is an p-dimensional *ellipsoid* with center $\mathbf{w}_0$.

The set of points

$$E = \{\mathbf{w} \mid (\mathbf{w} - \mathbf{w}_0)'\mathbf{Q}^{-1}(\mathbf{w} - \mathbf{w}_0) \leq K^2\} \qquad (1)$$

consists of the ellipsoid and the points inside it.

This handout explores different ways of defining and describing ellipsoids and their shapes.

An ellipse is a two dimensional ellipsoid. You can check that $q^{11} = q_{22}/D$, $q^{12} = -q_{12}/D$, and $q^{22} = q_{11}/D$, where $D = q_{11}q_{22} - q_{12}^2$.

You can explicitly compute w_2 as a function of w_1 on the boundary as

$$w_2 = w_{20} + [q_{12}(w_1-w_{10}) \pm \sqrt{D}\sqrt{\{q_{11}K^2 - (w_1-w_{10})^2\}}]/q_{11}, \qquad (2)$$

where the + and - signs go with the upper and lower halves of the ellipse. This is one way to find points on the boundary if you want to draw an ellipse. If you change every 1 subscript to 2 and every 2 subscript to 1 you get a similar equation that expresses w_1 in terms of w_2.

Incidentally, because you can't take the square root of a negative number, w_2 is defined only when $q_{11}K^2 - (w_1-w_{10})^2 \geq 0$, that is, when $(w_1-w_{10})^2 \leq q_{11}K^2$. This implies that if $\mathbf{w} = [w_1,w_2]'$ is on or in E it must always be the case that

$$-K\sqrt{q_{11}} \leq w_1-w_{10} \leq K\sqrt{q_{11}}$$

But this means that

$$w_{01} - K\sqrt{q_{11}} \leq w_1 \leq w_{01} + K\sqrt{q_{11}}.$$

Similarly, reversing the roles of w_1 and w_2, if $\mathbf{w} = [w_1,w_2]'$ is on or in E,

$$w_{02} - K\sqrt{q_{22}} \leq w_2 \leq w_{02} + K\sqrt{q_{22}}$$

In p dimensions, there are similar inequalities for *every* component w_i of $\mathbf{w}$:

$$w_{0i} - K\sqrt{q_{ii}} \leq w_i \leq w_{0i} + K\sqrt{q_{ii}},\ i = 1,\ldots,p \qquad (3)$$

These are particular cases of an important, more general family of inequalities:

For *any* p-vector $\boldsymbol{\ell} \neq \mathbf{0}$ and *any* $\mathbf{w}$ in E,

$$\boldsymbol{\ell}'\mathbf{w}_0 - K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \leq \boldsymbol{\ell}'\mathbf{w} \leq \boldsymbol{\ell}'\mathbf{w}_0 + K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \qquad (4)$$

When $\boldsymbol{\ell} = \mathbf{e}_i = [0\ 0\ \ldots\ 0\ 1\ 0\ \ldots]'$, where the 1 is in the i^{th} position, $\boldsymbol{\ell}'\mathbf{w} = w_i$, $\boldsymbol{\ell}'\mathbf{w}_0 = w_{0i}$ and $\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell} = \mathbf{q}_{ii}$ and equation (4) reduces to equation (3). Another way to express (4) is

$$\left|\boldsymbol{\ell}'\mathbf{w} - \boldsymbol{\ell}'\mathbf{w}_0\right| = \left|\boldsymbol{\ell}'(\mathbf{w} - \mathbf{w}_0)\right| \leq K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \qquad (4')$$

Eq. (4) and (4') apply to every $\mathbf{w}$ in E, and any p-dimensional $\boldsymbol{\ell}$.

Just as important is the converse:

For any $\mathbf{w}$ that is *not* in E, there is at least one $\boldsymbol{\ell} \neq \mathbf{0}$ for which eq. (4) is *not* satisfied. This means means that another *definition* for E is the set

$$E = \{\mathbf{w} \mid \boldsymbol{\ell}'\mathbf{w}_0 - K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \leq \boldsymbol{\ell}'\mathbf{w} \leq \boldsymbol{\ell}'\mathbf{w}_0 + K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \text{ for } \textit{every } \boldsymbol{\ell} \neq \mathbf{0}\} \quad (5)$$

Note that equation (5) involves $\mathbf{Q}$ itself and not $\mathbf{Q}^{-1}$.

For a point $\mathbf{w}$ to be on the boundary of E, there must be a $\boldsymbol{\ell} \neq \mathbf{0}$ such that

$$\boldsymbol{\ell}'\mathbf{w} = \boldsymbol{\ell}'\mathbf{w}_0 + K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \text{ or } \boldsymbol{\ell}'\mathbf{w} = \boldsymbol{\ell}'\mathbf{w}_0 - K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell}}.$$

In fact, such points are

$$\mathbf{w} = \mathbf{w}_0 + K\mathbf{Q}\boldsymbol{\ell}/\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})} \text{ or } \mathbf{w} = \mathbf{w}_0 - K\mathbf{Q}\boldsymbol{\ell}/\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}$$

For given $\boldsymbol{\ell}$, two sets

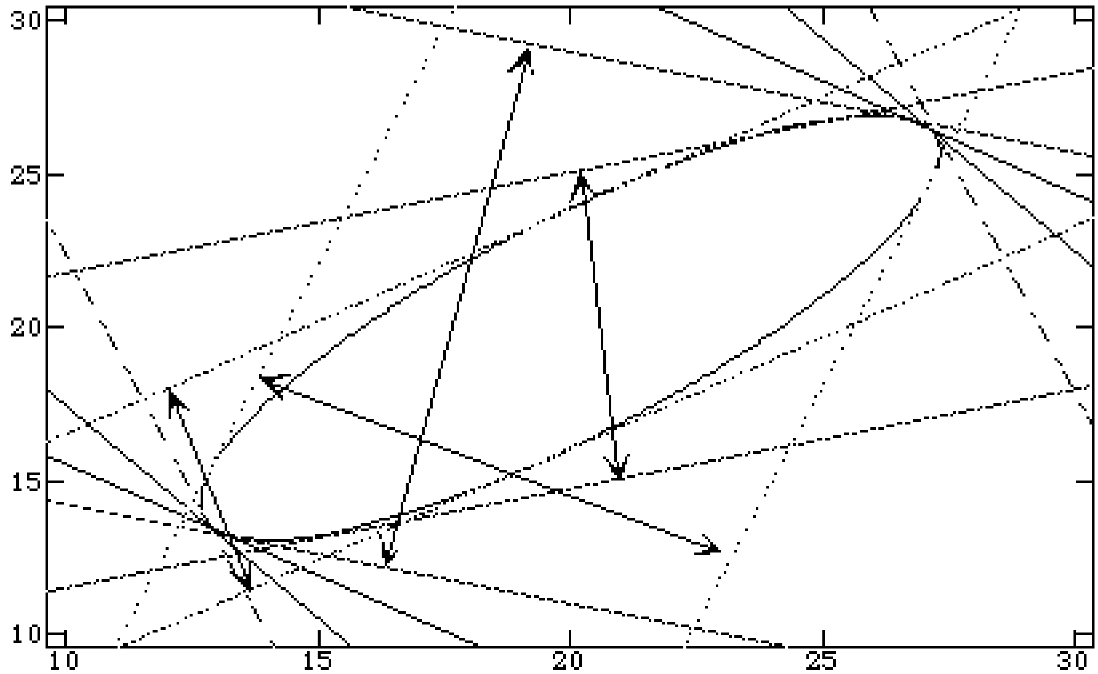
$$\{\mathbf{w} \mid \boldsymbol{\ell}'\mathbf{w} = \boldsymbol{\ell}'\mathbf{w}_0 - K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}\} \text{ and } \{\mathbf{w} \mid \boldsymbol{\ell}'\mathbf{w} = \boldsymbol{\ell}'\mathbf{w}_0 + K\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}\}$$

are parallel lines ($p = 2$) or planes ($p = 3$) or hyper-planes ($p > 3$) that are tangent to E. That is, they touch E only at the two boundary points $\mathbf{w}_0 \pm K\mathbf{Q}\boldsymbol{\ell}/\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}$ on opposite sides of E. These are the only two points on the tangent lines or planes that are actually in E. The rest of these tangent lines or planes are outside E.

If you consider $\boldsymbol{\ell}$ to be a vector from the origin ($\mathbf{w} = \mathbf{0}$), it determines a direction and these tangent lines or planes are perpendicular to that direction.

What equation 5 says is that an ellipsoid is defined as the set of all points between *every* pair of parallel tangent lines or planes. Thinking geometrically, this makes intuitive sense.

Here is an example of an ellipse centered at (20,20) with a number of pairs of parallel tangent lines drawn. The double headed arrows indicate some of the pairs of tangent lines



Because there is a tangent line or plane at every point on the boundary to E, every point $\mathbf{w}$ on the boundary can be expressed as $\mathbf{w} = \mathbf{w}_0 + K\mathbf{Q}\boldsymbol{\ell}/\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}$ for some $\boldsymbol{\ell}$. In fact, since $\mathbf{Q}\boldsymbol{\ell}/\sqrt{(\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell})}$ is unchanged if we replace $\boldsymbol{\ell}$ by the unit vector $\mathbf{u} = \boldsymbol{\ell}/\|\boldsymbol{\ell}\|$, every point $\mathbf{w}$ on the boundary can be expressed as $\mathbf{w} = \mathbf{w}_0 + K\mathbf{Qu}/\sqrt{(\mathbf{u}'\mathbf{Qu})}$ for some vector $\mathbf{u}$ with $\|\mathbf{u}\| = 1$.

When $p = 2$, by varying $\varnothing$, $0 \leq \varnothing < 2\pi$, you can get all unit vectors as $\mathbf{u} = [\cos\varnothing, \sin\varnothing]'$. Thus you can compute coordinates of points on an ellipse by

$$w_1 = w_{01} + K\,(q_{11}\cos\varnothing + q_{12}\sin\varnothing)/\sqrt{G(\varnothing)},$$
$$w_2 = w_{02} + K\,(q_{12}\sin\varnothing + q_{22}\cos\varnothing)/\sqrt{G(\varnothing)},$$

where $\varnothing$ varies between $-\pi$ to π or 0 to 2π and

$$G(\varnothing) = q_{11}\cos^2\varnothing + 2q_{12}\cos\varnothing\sin\varnothing + q_{22}\sin^2\varnothing$$
$$= (q_{11} + q_{22})/2 + [(q_{11} - q_{22})/2]\times\cos 2\varnothing + q_{12}\sin 2\varnothing.$$

This provides another way to compute points for plotting an ellipse, quite easy to do in MacAnova.

Suppose $\mathbf{T}$ is a "square root" of $\mathbf{Q}$, that is a p by p matrix satisfying $\mathbf{T}'\mathbf{T} = \mathbf{Q}$. Then $\mathbf{Q}^{-1} = \mathbf{T}^{-1}(\mathbf{T}')^{-1}$. Then, every point of the form $\mathbf{w} = \mathbf{w}_0 + K\mathbf{T}'\mathbf{u}$ where $\mathbf{u}$ is a unit vector ($\|\mathbf{u}\| = 1$) satisfies

$$(\mathbf{w} - \mathbf{w}_0)'\mathbf{Q}^{-1}(\mathbf{w} - \mathbf{w}_0)' = K^2\mathbf{u}'\mathbf{T}(\mathbf{T}^{-1}(\mathbf{T}')^{-1})\mathbf{T}'\mathbf{u} = K^2\mathbf{u}'\mathbf{u} = K^2,$$

and is therefore on the boundary of the ellipsoid. Another way to describe $\mathbf{w}$ is to note that $\mathbf{w} - \mathbf{w}_0 = K\mathbf{T}'\mathbf{u}$. When p = 2, this provides still another way to compute points $\mathbf{w} = \mathbf{w}_0 + K\mathbf{T}'\mathbf{u}$ on an ellipse, where again you get all unit vectors as $\mathbf{u} = [\cos\varnothing, \sin\varnothing]'$.

There are many matrices satisfying $\mathbf{T}_L'\mathbf{T}_L = \mathbf{Q}$. In particular, $\mathbf{T}_L$ can be chosen to be upper triangular ($p_{ij} = 0$ for $i > j$) with positive diagonal ($p_{jj} > 0$, $j = 1,\ldots,p$). In this case the representation $\mathbf{T}_L'\mathbf{T}_L = \mathbf{Q}$ is the well known *Cholesky* decomposition of $\mathbf{Q}$. It can be computed in MacAnova by `tL <- cholesky(q)` or `tL <- matsqrt(q)`.

When p = 2, the upper triangular square root of $\mathbf{Q}$ is

$$\mathbf{T}_L = \begin{bmatrix} \sqrt{q_{11}} & q_{12}/\sqrt{q_{11}} \\ & \\ 0 & \sqrt{(D/q_{11})} \end{bmatrix}, \; D = q_{11}q_{22} - q_{12}^2.$$

This provides another way to compute points on the ellipse as

$$w_1 = w_{01} + \sqrt{q_{11}}\cos\varnothing, \quad w_2 = w_{01} + (q_{12}\cos\varnothing + \sqrt{D}\sin\varnothing)/\sqrt{q_{11}}$$

and letting $\varnothing$ vary over $-\pi$ to π or 0 to 2π. This is also easy to do in MacAnova.

Another important square root of $\mathbf{Q}$ is the *symmetric* square root $\mathbf{T}_S$ which you can compute in MacAnova by `tS <- matsqrt(q, symmetric:T)`. Briefly, $\mathbf{T}_S$ has the same eigenvectors as $\mathbf{Q}$, but its eigenvalues are the square roots of the eigenvalues of $\mathbf{Q}$.

When p = 2, there is no simple expression for $\mathbf{T}_S$ except in terms of the eigenvalues and vectors of $\mathbf{Q}$.

The distance from the center of the ellipse $\mathbf{w}_0$ to the point $\mathbf{w} = \mathbf{w}_0 + K\mathbf{T}_S'\mathbf{u}$ on the boundary and the opposite point $\mathbf{w} = \mathbf{w}_0 + K\mathbf{T}_S'(-\mathbf{u})$ is

$$\|\mathbf{w} - \mathbf{w}_0\| = \sqrt{\{(\mathbf{w} - \mathbf{w}_0)'(\mathbf{w} - \mathbf{w}_0)\}} =$$
$$K\sqrt{\{\mathbf{u}'\mathbf{T}_S\mathbf{T}_S'\mathbf{u}\}} = K\sqrt{\{\mathbf{u}'\mathbf{T}_S\mathbf{T}_S\mathbf{u}\}} = K\sqrt{\{\mathbf{u}'\mathbf{Q}\mathbf{u}\}}.$$

Eigenvectors and eigenvectors have important properties related to the extremes of $\mathbf{u}'\mathbf{Q}\mathbf{u}$ for unit vectors $\mathbf{u}$. In particular, recalling that by convention the eigenvalues are numbered in decreasing order,

$$\max_{\|\mathbf{u}\|=1}\mathbf{u}'\mathbf{Q}\mathbf{u} = \mathbf{u}_1'\mathbf{Q}\mathbf{u}_1 = \lambda_1 \text{ and } \min_{\|\mathbf{u}\|=1}\mathbf{u}'\mathbf{Q}\mathbf{u} = \mathbf{u}_p'\mathbf{Q}\mathbf{u}_p = \lambda_p \qquad (6)$$

Now you can express every point $\mathbf{w}$ on the boundary as $\mathbf{w} = \mathbf{w}_0 + K\mathbf{T}_S'\mathbf{u} = \mathbf{w}_0 + K\mathbf{T}_S\mathbf{u}$ for some unit vector $\mathbf{u}$, and we have just seen that the distance of such a point from $\mathbf{w}_0$ is $K\sqrt{\{\mathbf{u}'\mathbf{Q}\mathbf{u}\}}$. Thus the points on the boundary that are furthest from $\mathbf{w}_0$ are $\mathbf{w}_0 \pm K\mathbf{T}_S\mathbf{u}_1 = \mathbf{w}_0 \pm K\sqrt{\lambda_1}\mathbf{u}_1$ at distance $K\sqrt{\lambda_1}$, and the closest points to the center are $\mathbf{w}_0 \pm K\sqrt{\lambda_p}\mathbf{u}_p$ at distance $K\sqrt{\lambda_p}$ from $\mathbf{w}_0$. ($\mathbf{T}_S\mathbf{u}_i = \sqrt{\lambda_i}\mathbf{u}_i$ because $\mathbf{u}_i$ is an eigenvector of $\mathbf{T}_S$ with eigenvalue $\sqrt{\lambda_i}$.)

The p lines connecting the opposite points $\mathbf{w}_0 - K\sqrt{\lambda_i}\mathbf{u}_i$ = and $\mathbf{w}_0 + K\sqrt{\lambda_i}\mathbf{u}_i$, i = 1,...,p are the *principal axes* of the ellipsoid. Each goes through the center $\mathbf{w}_0$ and has length $2K\sqrt{\lambda_i}$. The first such line is the *major axis* and its length is the longest distance between opposite points on the surface of the ellipsoid. The p^{th} line is the *minor axis* of the ellipsoid and its length is the shortest distance between opposite points. The principal axes are all mutually orthogonal (perpendicular).

Let $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be positive definite and E_1 and E_2 be ellipsoids centered at $\mathbf{w}_{01}$ and $\mathbf{w}_{02}$ defined by

$$E_i = \{\mathbf{w} \mid (\mathbf{w} - \mathbf{w}_{0i})'\mathbf{Q}_i^{-1}(\mathbf{w} - \mathbf{w}_{0i}) \leq K_i^2\},\ i = 1, 2,$$

For the moment assume, $K_1 = K_2$. Then when $\mathbf{Q}_1$ and $\mathbf{Q}_2$ have the *same* eigenvalues, E_1 and E_2 have the same size and shape, in the sense that E_2 can be obtained from E_1 by moving it so that $\mathbf{w}_{01}$ coincides with $\mathbf{w}_{02}$ and then possibly rotating it so that the corresponding principal axes of the two ellipsoids are the same. Rotation is needed to make them coincide if and only if the eigenvectors of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ differ. Thus when the eigenvectors differ, the ellipses have different *orientations*. If the eigenvalues are the same, but $K_1 > K_2$, E_1 and E_2 have the same shape, but different sizes, with E_1 larger than E_2.

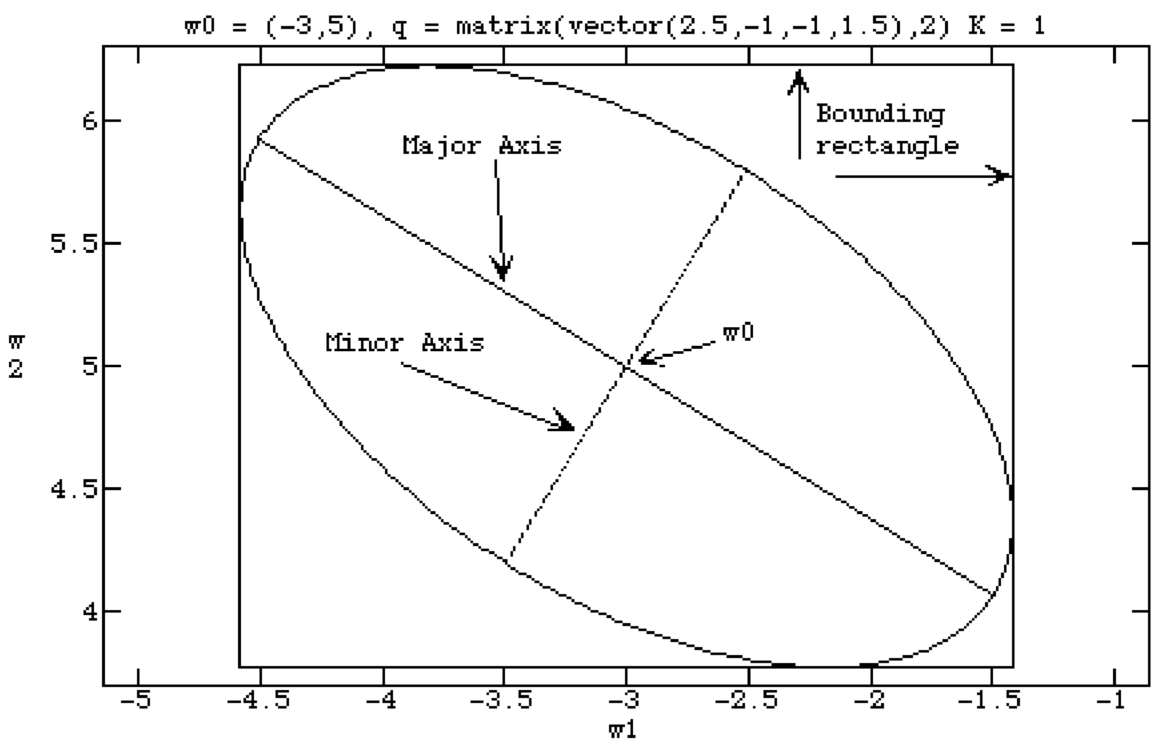
When the eigenvalues $\{\lambda_{i1}\}$ and $\{\lambda_{i2}\}$ of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ differ, then E_1 and E_2 have different shapes except in the case when the eigenvalues are proportional, that is $\lambda_{i2} = c\lambda_{i1}$ for some $c > 0$. It this case, the ellipsoids have the same shapes, but different sizes unless $K_2 = cK_1$.

Thus the eigenvalues of **Q**, or, more accurately, their ratios, are descriptive of the shapes of ellipsoids, and the eigenvectors are descriptive of their orientation.

Below is a graph of an ellipse centered at $\mathbf{w}_0$ = [-3, 5]' with

$$\mathbf{Q} = \begin{bmatrix} 2.5 & -1 \\ -1 & 1.5 \end{bmatrix},$$

with eigenvectors $\mathbf{u}_1$ = [0.85065, -0.52573]' and $\mathbf{u}_2$ = [0.52573, 0.85065]', at angles 121.72° and 31.72°, respectively, with the w_1 axis. The eigenvalues of **Q** are $\lambda_1 = 3.118 = 1.766^2$ and $\lambda_2 = 0.8820 = 0.9391^2$ and hence the lengths of the major and minor axes are $2\sqrt{\lambda_1} = 3.532$ and 1.878, respectively. If you rotate the ellipse and move it to a different position, the shape would be the same. It would be described by different **Q** and $\mathbf{w}_0$, say $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{w}}_0$, but the eigenvalues of $\tilde{\mathbf{Q}}$ would always be the same as those of **Q**. The eigenvectors of $\tilde{\mathbf{Q}}$ would define the directions of the principal axes or the rotated ellipse.



Application to statistics

Many confidence regions in statistics are ellipsoids that may be expressed in the form

$$R = \{\boldsymbol{\theta} \mid (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})'(\hat{V}[\hat{\boldsymbol{\theta}}])^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq K^2\},$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_q]'$ is a vector of q parameters, $\hat{\boldsymbol{\theta}}$ is an estimator for $\boldsymbol{\theta}$, $\hat{V}[\hat{\boldsymbol{\theta}}]$ is the estimated variance/covariance matrix of $\hat{\boldsymbol{\theta}}$, and K^2 is a critical value. In large samples, K^2 is often $\chi_{q,1-\alpha}{}^2$. This is an ellipsoid of the form of (1) with $\mathbf{Q} = \hat{V}[\hat{\boldsymbol{\theta}}]$ centered at $\mathbf{w}_0 = \hat{\boldsymbol{\theta}}$.

Here are two possibly familiar examples:

(i) A confidence region for the vector of least squares regression coefficients $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_k]'$ in a univariate multiple regression of y on $\mathbf{x} = [1, x_1, \ldots, x_k]'$ is

$$\begin{aligned} R &= \{\boldsymbol{\beta} \mid (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'\mathbf{X}'\mathbf{X}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})/s^2 \leq (k+1)\, F_{k+1,n-k-1}(1-\alpha)\} \\ &= \{\boldsymbol{\beta} \mid (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'\mathbf{X}'\mathbf{X}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \leq s^2(k+1)\, F_{k+1,n-k-1}(1-\alpha)\}, \end{aligned}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]'$ is n by k+1 and s^2 is the residual mean square error. Recall that $\hat{V}[\hat{\boldsymbol{\beta}}] = s^2(\mathbf{X}'\mathbf{X})^{-1}$. In terms of the definition in eq. (1), $\mathbf{Q} = (\mathbf{X}'\mathbf{X})^{-1}$.

(ii) A confidence region for the mean vector of a p-dimensional multivariate normal population mean $\boldsymbol{\mu}$ based on a random sample $[\mathbf{y}_1, \ldots, \mathbf{y}_n]'$ is

$$R = \{\boldsymbol{\mu} \mid (\boldsymbol{\mu} - \overline{\mathbf{y}})'[(1/n)\mathbf{S}]^{-1}(\boldsymbol{\mu} - \overline{\mathbf{y}}) \leq (p(n-1)/(n-p))F_{p,n-p}(1-\alpha)\}.$$

The preceding discussion provides a basis for describing such confidence intervals. Here $\mathbf{Q} = (1/n)\mathbf{S}$].

The shapes and orientations of these regions are determined by the eigenvalues and eigenvectors of $\hat{V}[\hat{\boldsymbol{\theta}}]$. Moreover, the quantity corresponding to $\boldsymbol{\ell}'\mathbf{Q}\boldsymbol{\ell}$ is $\boldsymbol{\ell}'\hat{V}[\hat{\theta}]\boldsymbol{\ell} = \hat{V}[\boldsymbol{\ell}'\hat{\theta}]$, the variance of a linear combination of the parameters. Thus, for any $\boldsymbol{\theta}$ in *R*, $\boldsymbol{\ell}'\boldsymbol{\theta}$ must satisfy

$$\boldsymbol{\ell}'\hat{\boldsymbol{\theta}} - K\sqrt{\hat{V}[\boldsymbol{\ell}'\hat{\boldsymbol{\theta}}]} \leq \boldsymbol{\ell}'\boldsymbol{\theta} \leq \boldsymbol{\ell}'\hat{\boldsymbol{\theta}} + K\sqrt{\hat{V}[\boldsymbol{\ell}'\hat{\boldsymbol{\theta}}]},$$

and $\hat{\boldsymbol{\theta}}$ is in *R* only if this is satisfied for every $\boldsymbol{\ell}$. Thus, the probability that these inequalities are satisfied for *all* $\boldsymbol{\ell}$ is $1 - \alpha$ = P(*R* contains true $\boldsymbol{\theta}$). They therefore constitute a set of *simultaneous* confidence intervals for *all* linear combinations of the elements of $\boldsymbol{\theta}$. It is noteworthy that every interval is of the same *form* as an ordinary confidence interval, for which K might be $z_{1-\alpha/2}$ or $t_{f,1-\alpha/2}$.

In particular, when $\boldsymbol{\ell} = \mathbf{e}_i$, $\boldsymbol{\ell}'\hat{V}[\hat{\boldsymbol{\theta}}]\boldsymbol{\ell} = \hat{V}[\hat{\theta}_i]$ is the estimated variance of the estimate of the i^{th} parameter, and the minimum and maximum values of θ_i for any $\boldsymbol{\theta}_i$ in *R* are $\hat{\theta}_i - K\sqrt{\hat{V}[\hat{\theta}_i]}$ and $\hat{\theta}_i + K\sqrt{\hat{V}[\hat{\theta}_i]}$. These intervals are *very* conservative simultaneous confidence intervals for the individual elements of $\boldsymbol{\theta}$. If you need simultaneous intervals for just $\theta_1, \ldots, \theta_q$, those based on the Bonferroni inequality using $K' = z_{1-(\alpha/q)/2}$ or $t_{f,1-(\alpha/q)/2}$ instead of K will be much shorter.

For example (i), the eigenvectors of $\hat{V}[\hat{\boldsymbol{\theta}}]$ are fixed (not random) and are just the eigenvectors of $\mathbf{Q} = (\mathbf{X}'\mathbf{X})^{-1}$; the eigenvalues are $s^2\lambda_i$, where $\lambda_1,\ldots,\lambda_{k+1}$ are the eigenvalues of $\mathbf{Q}$. Thus the orientation and shape of *R* is depends only on the independent variables; its center is $\hat{\boldsymbol{\beta}}$ and its size is proportional to s.

The minimum and maximum values for β_i for any $\boldsymbol{\beta}$ in the confidence region *R* are $\hat{\beta}_i \pm K\sqrt{\hat{V}[\hat{\beta}_i]} = \hat{\beta}_i \pm Ks\sqrt{c_{ii}}$, where $K = \sqrt{\{(k+1)\,F_{k+1,n-k-1}(1-\alpha)\}}$ and c_{ii} is the i^{th} diagonal element of $\mathbf{Q}$. They constitute very conservative simultaneous confidence limits for all the individual elements of $\boldsymbol{\beta}$. The Bonferronized limits re preferable to these. They are computed the same way using

$$K' = t_{n-k-1,1-(\alpha/(k+1))/2}$$

instead of K. Simultaneous limits for *all* linear combinations are

$$\boldsymbol{\ell}'\boldsymbol{\beta} = \boldsymbol{\ell}'\hat{\boldsymbol{\beta}} \pm Ks\sqrt{\{\boldsymbol{\ell}'(\mathbf{X}'\mathbf{X})^{-1}\,\boldsymbol{\ell}\}},$$

the so-called Scheffe confidence limits.

For example (ii), the eigenvectors are the same as the eigenvectors of the random matrix $\mathbf{S}$ and the eigenvalues are 1/n times the eigenvalues of $\mathbf{S}$. Thus both the orientation and shape of the confidence ellipsoid depends on the sample variance matrix and will vary from sample to sample.

The minimum and maximum values for μ_i for any $\boldsymbol{\mu}$ in confidence region *R* are $\overline{y_i} \pm K\sqrt{\hat{V}[\overline{y_i}]} = \overline{y_i} \pm K\sqrt{(s_{ii}/n)}$, where $K = \sqrt{\{p(n-1)F_{1-\alpha,p,n-p}/(n-p)\}}$. These are very conservative simultaneous confidence limits for $\mu_1, \ldots, \mu_p$. Bonferroni based limits using $K' = t_{n-1,1-(\alpha/p)/2}$ are much shorter.

Simultaneous confidence intervals for all linear combinations $\boldsymbol{\ell}'\boldsymbol{\mu}$ are given by $\boldsymbol{\ell}'\boldsymbol{\mu} \pm K\,\sqrt{\{\boldsymbol{\ell}'\mathbf{S}\boldsymbol{\ell}'/n\}}$. Only if you are interested in a very large number of different $\boldsymbol{\beta}$'s will these intervals be shorter than Bonferronized limits.