# ELLIPSES

**Problem:** Find the points on the locus

$$Q(x,y) = 865x^2 - 294xy + 585y^2 = 1450$$

closest to, and farthest from, the origin.

**Answer.**

This is a Lagrange multiplier problem: we want to extremize $f(x,y) = x^2 + y^2$ subject to the constraint $Q(x,y) = 1450$. To do this we look for points on the locus where the gradient of $f + \lambda Q$ is zero: this amounts to solving the system

$$\nabla(f + \lambda Q) = (2x + \lambda(1730x - 294y), 2y + \lambda(-294x + 1070y)) = (0,0)$$

(together with the constraint equation: that gives us three equations in three unknowns ($x, y$, and $\lambda$). But if we stare at this for a minute, we start to suspect that it might be easier to deal algebraically with the related problem, of finding the maximum of the function $Q$ on the unit circle $f(x,y) = 1$: that question requires us to find the points where the gradient of $Q + \lambda f$ vanishes – which is pretty much what we would get if we multiplied our original problem by $\lambda^{-1}$. After eliminating superfluous factors of two, the new problem gives us the system

$$\lambda x + 865x - 147y = 0 \ , \ \lambda y - 147x + 565y = 0 \ ,$$

which can be rewritten in matrix notation in the form

$$(A + \lambda \mathbf{1})\mathbf{x} = 0 \ ,$$

where $\mathbf{x}$ denotes the vector $(x,y)$, $\mathbf{1}$ denotes the two-by-two 'identity' matrix (with ones down the diagonal, and zero elsewhere), and

$$A = \left[ \begin{array}{cc} 865 & -147 \\ -147 & 585 \end{array} \right] .$$

Now it's useful to recall that square matrices are **invertible** (ie, have **inverse** matrices, in the sense the the matrix product of a matrix and its inverse equals the identity matrix) if and only if its determinant is nonzero. [This is related to the geometric interpretation of determinants as volumes: if the determinant vanishes, the linear transformation defined by the (square) matrix squashes a rectangle flat.] So if the determinant of the matrix $A + \lambda\mathbf{1}$ is **not** zero, there can be no nontrivial solutions to our system of equations – because we can then multiply our equation on the left by the inverse matrix $(A + \lambda\mathbf{1})^{-1}$, to obtain (only) the trivial solution

$$(A + \lambda\mathbf{1})^{-1}(A + \lambda\mathbf{1})\mathbf{x} = \mathbf{1} \cdot \mathbf{x} = \mathbf{x} = 0 \ .$$

Thus in order for a nontrivial solution to exist, $\lambda$ must satisfy the quadratic equation $\det(A + \lambda\mathbf{1}) = 0$, ie

$$\det \left[ \begin{array}{cc} 865 + \lambda & -147 \\ -147 & 585 + \lambda \end{array} \right] = (\lambda + 865)(\lambda + 585) - (-147)^2 = 0 \ ,$$

which multiplies out to

$$\lambda^2 + 1450\lambda + [(865 \times 585 - 147^2) = 484,416] = 0 \ .$$

According to the quadratic formula, then,

$$\lambda = \tfrac{1}{2}[-1450 \pm \sqrt{(1450^2 - 4 \times 484,416)}] \ ;$$

but

$$1450^2 - 4 \times 484,416 = 2,102,500 - 1,937,664 = 164,836 = 406^2 \ ,$$

so

$$\lambda = \tfrac{1}{2}[-1450 \pm 406] = -522 \ (= -9 \cdot 58) \text{ or } -928 \ (= -16 \cdot 58) \ .$$

Substituting the first of these values into our original matrix equation gives us

$$\begin{bmatrix} 865 - 522 & -147 \\ -147 & 585 - 522 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \ ,$$

but the matrix factors as

$$\begin{bmatrix} 343 & -147 \\ -147 & 63 \end{bmatrix} = \begin{bmatrix} 7 \cdot 49 & -3 \cdot 49 \\ -7 \cdot 21 & 3 \cdot 21 \end{bmatrix}$$

and thus kills the vector $(3, 7)$ – as well as its normalized multiple

$$\mathbf{x}_+ = \frac{1}{\sqrt{58}}(3, 7) \ .$$

Similarly: substituting in the second root yields the matrix

$$\begin{bmatrix} 865 - 928 & -147 \\ -147 & 585 - 928 \end{bmatrix} = \begin{bmatrix} -63 & -147 \\ -147 & -343 \end{bmatrix} \ ,$$

which admits a similar factorization, and consequently kills the normalized vector

$$\mathbf{x}_- = \frac{1}{\sqrt{58}}(-7, 3) \ .$$

We can use these to answer the original question, about points on the curve $Q = 1450$ at greatest and least distance from the origin, by noting that the function $Q$ is quadratic, in the sense that for any real number $t$, we have

$$Q(tx, ty) = t^2 Q(x, y) \ ,$$

The points on the locus $Q = 1450$ where $f$ is greatest are just multiples $t\mathbf{x}_\pm$ of the points on the unit circle where $Q$ is greatest, by the argument above (about inverting $\lambda$); so all we need to do is find the right scaling factor $t$. It's not hard to calculate that

$$Q(\mathbf{x}_+) = 58 \cdot 9 \ , \ Q(\mathbf{x}_-) = 58 \cdot 16 \ ,$$

from which it follows easily that $\frac{5}{3}\mathbf{x}_+$ and $\frac{5}{4}\mathbf{x}_-$ lie on the locus $Q = 1450 = 25 \cdot 58$: they are the extreme points we sought. Note by the way that these vectors (like $\mathbf{x}_+$ and $\mathbf{x}_-$) are **perpendicular**: their dot product is

$$\frac{25}{12 \cdot 58} \ (3 \cdot (-7) + 7 \cdot 3) = 0 \ .$$

In other words: the locus $Q = 1450$ is an ellipse, with the first vector above as the semimajor, and the second vector the semiminor, axes. It can be obtained from the standard ellipse

$$9X^2 + 16Y^2 = 25$$

by applying the **rotation matrix**

$$[\mathbf{x}_+ \mathbf{x}_-] := \frac{1}{\sqrt{58}} \begin{bmatrix} 3 & -7 \\ 7 & 3 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} := R(\theta)$$

through the angle satisying $\tan\theta = \sin\theta/\cos\theta = 7/3$. [Note that

$$R(\theta)\cdot \mathbf{e}_1 = \mathbf{x}_+,\ R(\theta)\cdot \mathbf{e}_2 = \mathbf{x}_- \ ,$$

where $\mathbf{e}_1 = (1,0), \mathbf{e}_2 = (0,1)$ are the standard unit vectors.] In fact

$$Q(x,y) = 9(3x+7y)^2 + 16(-7x+3y)^2 \ ;$$

writing this out gives

$$9(9x^2 + 42xy + 49y^2) + 16(49x^2 - 42xy + 9y^2) \ ,$$

which equals

$$[9\cdot 9 + 16\cdot 49 = 865]x^2 + [42\cdot(9-16) = -294]xy + [9\cdot 49 + 16\cdot 9 = 585]y^2 \ .$$

**2** For clarity, here is another example, this time with smaller numbers:

**Problem:** Find the **principal axes** (ie the semimajor and semiminor axes) for the ellipse $Q(x,y) = 23x^2 + 14xy + 23y^2 = 17$.

**Solution:** We need to find the **eigenvectors** of the matrix

$$\left[\begin{array}{cc} 23 & 7 \\ 7 & 23 \end{array}\right] = B \ ;$$

these are the (nontrivial) vectors $\mathbf{v}_\pm$ satisfying the **eigenvalue** equation $(B + \lambda_\pm)\mathbf{v}_\pm = 0$. [**Eigen** comes from German, where it signifies something like 'proper' or 'characteristic'. It has become standard in mathematics (and in quantum mechanics) in this and related contexts.] Thus we need to solve the quadratic equation

$$\det \left[\begin{array}{cc} 23+\lambda & 7 \\ 7 & 23+\lambda \end{array}\right] = (23+\lambda)^2 - 7^2 = \lambda^2 + 46\lambda + [529 - 49 = 480] = 0 \ .$$

This has solutions

$$\lambda_\pm = \tfrac{1}{2}[-46 \pm \surd(46^2 - 4\cdot 480)] \ ;$$

but the term under the square root sign equals $2116 - 1920 = 196 = 13^2$, so $\lambda_+ = -16$ and $\lambda_- = -30$. It follows that

$$B + \lambda_+ = \left[\begin{array}{cc} 23-16 & 7 \\ 7 & 23-16 \end{array}\right] = \left[\begin{array}{cc} 7 & 7 \\ 7 & 7 \end{array}\right]$$

which kills the normalized vector $\mathbf{v}_+ = \frac{1}{\sqrt{2}}(1,-1)$, while

$$B + \lambda_- = \left[\begin{array}{cc} 23-30 & 7 \\ 7 & 23-30 \end{array}\right] = \left[\begin{array}{cc} -7 & 7 \\ 7 & -7 \end{array}\right]$$

kills the vector $\mathbf{v}_- = \frac{1}{\sqrt{2}}(1,1)$. In this case the rotation matrix is

$$R(\phi) = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} 1 & 1 \\ -1 & 1 \end{array}\right] \ ,$$

defined by a 45-degree rotation (since $\tan\phi = 1$), and the original equation can be rewritten as

$$\lambda_+(\mathbf{v}_+\cdot\mathbf{x})^2 + \lambda_-(\mathbf{v}_-\cdot\mathbf{x})^2 = 8(x-y)^2 + 15(x+y)^2 \ .$$

In general the normalized eigenvectors satisfy $Q(\mathbf{v}_\pm) = -\lambda_\pm$, so by rescaling (as in the previous problem) we find that the principal axes of the ellipse $Q = 17$ are defined by the vectors

$$\frac{\sqrt{34}}{8}\,(1,1) \text{ and } \frac{\sqrt{1120}}{60}\,(1,-1) \ .$$

**3** In general, the $n$-dimensional Lagrange multiplier problem for a quadratic function

$$\mathbf{x} \mapsto (A\mathbf{x}) \cdot \mathbf{x} : \mathbb{R}^n \to \mathbb{R}$$

defined by a symmetric $n \times n$ matrix $A$ will have $n$ nontrivial (normalized, mutually orthogonal) eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ satisfying the eigenvalue equation $A\mathbf{v}_k = \lambda_k \mathbf{v}_k$, and the formula above generalizes to

$$Q(\mathbf{x}) = \sum_{k=1}^{k=n} \lambda_k (\mathbf{v}_k \cdot \mathbf{x})^2 \ .$$

Astronomers use this, for example, to study elliptical galaxies (which, being three-dimensional objects, have **three** principal axes).

I don't want to give the (false) impression, that the quadratic equations for such problems always work out so neatly; they don't. The problems above were cooked up using the equation

$$A(ax + by)^2 + B(-bx + ay)^2 = Ex^2 + 2Fxy + Gy^2 \ ,$$

where

$$E = a^2 A + b^2 B, \ F = ab(A - B), \ \text{and } G = b^2 A + a^2 B \ ;$$

for example in problem 1, $A = 9, \ B = 16, \ a = 3$, and $b = 7$. If you chase through the quadratic formula in this generality, you find that for problems of this sort,

$$\lambda_\pm = -(a^2 + b^2)\{A \text{ or } B\} \ .$$

**4** The condition that the matrix $A$ be symmetric is important. The symmetry condition [that the coefficient $A_{ik}$ of the matrix equals the coefficient $A_{ki}$, with the order of indices reversed] implies that for any two vectors $\mathbf{v}$ and $\mathbf{v}'$, we have

$$(A\mathbf{v}') \cdot \mathbf{v} = \sum_{i,k=1}^{i,k=n} A_{ik} v'_k v_i = \sum_{i,k=1}^{i,k=n} A_{ki} v_i v'_k = (A\mathbf{v}) \cdot \mathbf{v}' \ .$$

But now if $\mathbf{v}$ and $\mathbf{v}'$ are eigenvectors of a symmetric matrix $A$, with associated eigenvalues $\lambda$ and $\lambda'$ which are **distinct**, ie $\lambda \neq \lambda'$, then $\mathbf{v}$ and $\mathbf{v}'$ must be orthogonal: for

$$A\mathbf{v} = \lambda\mathbf{v}, \ A\mathbf{v}' = \lambda'\mathbf{v}' \ ,$$

so on one hand

$$(A\mathbf{v}') \cdot \mathbf{v} = (\lambda'\mathbf{v}') \cdot \mathbf{v} = \lambda'\mathbf{v}' \cdot \mathbf{v} \ ;$$

while on the other hand, the symmetry of $A$ implies that

$$(A\mathbf{v}') \cdot \mathbf{v} = (A\mathbf{v}) \cdot \mathbf{v}' = (\lambda\mathbf{v}) \cdot \mathbf{v}' = \lambda\mathbf{v} \cdot \mathbf{v}' \ .$$

Thus these two quantities are equal, ie

$$\lambda'\mathbf{v}' \cdot \mathbf{v} = \lambda\mathbf{v} \cdot \mathbf{v}' \ ;$$

but the dot product itself is symmetric [$\mathbf{v}' \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{v}'$] so

$$(\lambda' - \lambda)\mathbf{v}' \cdot \mathbf{v} = 0 \ .$$

But $\lambda$ and $\lambda'$ are distinct, so their difference is nonzero, and we conclude that $\mathbf{v}' \cdot \mathbf{v} = 0$ – which is to say that the eigenvectors $\mathbf{v}$ and $\mathbf{v}'$ are perpendicular.

This has applications in quantum mechanics: in that theory observable physical quantities are supposed to be represented by (things very much like) symmetric matrices, and the result of a measurement of the physical quantity in question is thought to be an eigenvalue of that matrix. A **state** of the physical system is interpreted as a vector, and to say that a measurement of a physical quantity $A$ in the state $\mathbf{v}$ yields the result $\lambda$ is interpreted as saying that the state $\mathbf{v}$ is an eigenvector of $A$, with eigenvalue $\lambda$: in other words,

$$A\mathbf{v} = \lambda\mathbf{v} .$$

The fact that eigenvectors corresponding to distinct eigenvalues are orthogonal is a kind of quantum-mechanical analog of the law of the excluded middle in logic: there is a certain amount of indeterminacy in quantum mechanics - an experiment might yield $\lambda$ for a measurement, or it might yield $\lambda'$ - but it can't yield both: the experiment yields a 'pure' state, in which the value of the measured quantity is well-defined.

This is a difficult and important notion, which lies at the heart of quantum mechanics, and it's really because I thought you might be interested in this, rather than because of the considerable intrinsic mathematical beauty of the theory of principal axes, that I have written up these notes.