5. **Properties of ellipsoids,** (p.14).

   (a) Draw the ellipsoid:
   $$ax^2 + by^2 = 1 \tag{9}$$
   What are the directions and lengths of the principal axes?

   (b) Draw the ellipsoid:
   $$2x^2 + 2xy + 2y^2 = 1 \tag{10}$$
   What are the directions and lengths of the principal axes?

   (c) Given an N-dimensional ellipsoid:
   $$x^T W x = 1 \tag{11}$$
   where W is a real, symmetric, positive definite matrix. What are the directions and lengths of the principal axes?

   (Hint: start with part (c).)

Solution for problem 5. (p.4) We begin with part (c), which is the general problem. Then we proceed to the specific cases, parts (a) and (b), which entail direct applications of the general solution.

(c) Let x be an arbitrary point on an N-dimensional ellipsoid centered at the origin and defined by:

$$x^T W x = 1 \tag{48}$$

where W is a real, symmetric, positive definite matrix.

We wish to characterize the principal axes of the ellipsoid. A point x is a principal axis of the ellipsoid if $\|x\|$ (the length of x) is an extremum (a stationay point) among the points on the ellipsoid. So, to find the principal axes we must find all the solutions of the following optimization problem:

$$\text{optimize } \|x\| \quad \text{subject to the constraint:} \quad x^T W x = 1 \tag{49}$$

The length of the vector x is:

$$\|x\| = \sqrt{x^T x} \tag{50}$$

We obtain the same extrema if, in (49), we optimize $\|x\|^2$ instead, since the quadratic function is monotonic. Consequently, our optimization problem can be expressed as:

$$\text{optimize } x^T x \quad \text{subject to the constraint:} \quad x^T W x = 1 \tag{51}$$

We use Lagrange optimization to find the desired extrema. Define the function:

$$H = x^T x + \lambda(1 - x^T W x) \tag{52}$$

The quantity in parentheses equals zero when the constraint ($x^T W x = 1$) is satisfied. Hence, optimizing H subject to the constraint is the same as optimizing $x^T x$ subject to the constraint. We optimize H by equating its derivative to zero:

$$0 = \frac{\partial H}{\partial x} = 2x - 2\lambda W x \tag{53}$$

This implies that any point x which is a constrained optimum of H is an eigenvector of W:

$$W x = \frac{1}{\lambda} x \tag{54}$$

Furthermore, $\frac{1}{\lambda}$ is the corresponding eigenvalue.

W is a real, symmetric, square matrix so it has a full complement of N orthogonal eigenvectors. Let us denote the orthonormal eigenvectors of W by $v_1, \ldots, v_N$, with corresponding eigenvalues $\mu_1, \ldots, \mu_N$:

$$W v_n = \mu_n v_n, \quad n = 1, \ldots, N, \quad v_n^T v_m = \delta_{mn} \tag{55}$$

where δ_{mn} is the Kronecker delta function.

Any principal axis x is an eigenvector of W, so it is proportional to one of the orthonormal eigenvectors:

$$x = c v_n \tag{56}$$

where c is a constant determined by the constraint:

$$x^T W x = 1 \quad \Longrightarrow \quad c^2 v_n^T W v_n = 1 \tag{57}$$

Using eq.(55), the righthand relation implies:

$$c = \pm \frac{1}{\sqrt{\mu_n}} \tag{58}$$

We now have a full characterization of the principal axes of an N-dimensional ellipsoid centered at the origin and defined by the matrix W. Its nth semi-axis is:

$$x = \frac{1}{\sqrt{\mu_n}} v_n \tag{59}$$

whose length is:

$$\sqrt{x^T x} = \frac{1}{\sqrt{\mu_n}} \tag{60}$$

We now proceed to solve the special cases in parts (a) and (b).

(a) The ellipse in eq.(9) can be represented in vector form as:

$$(x\ y)^T \underbrace{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}}_{W} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \tag{61}$$

which defines the matrix W.

The eigenvalues of W are the solutions of the characteristic equation:

$$0 = \begin{vmatrix} a-\lambda & 0 \\ 0 & b-\lambda \end{vmatrix} = (a-\lambda)(b-\lambda) \tag{62}$$

which implies that the eigenvalues are:

$$\mu_1 = a, \quad \mu_2 = b \tag{63}$$

Each eigenvector is the solution of:

$$W v_n = \mu_n v_n \tag{64}$$

which is two linearly dependent equations in the two unknowns of the vector v_n. The orthonormal solutions are found to be:

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{65}$$

So, the principal axes are aligned along the x- and y-axes. The length of the x-semi-axis is $1/\sqrt{a}$, while the length of the y-semi-axis is $1/\sqrt{b}$. More explicitly, the semi-axes are:

$$s_1 = \frac{1}{\sqrt{\mu_1}} v_1 = \frac{1}{\sqrt{a}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad s_2 = \frac{1}{\sqrt{\mu_2}} v_2 = \frac{1}{\sqrt{b}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{66}$$

(b) The ellipse in eq.(10) can be represented in vector form as:

$$(x\ y)^T \underbrace{\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}}_{W} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \tag{67}$$

which defines the matrix W.

The eigenvalues of W are the solutions of the characteristic equation:

$$0 = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 \tag{68}$$

which implies that the eigenvalues are:

$$\mu_1 = 1, \quad \mu_2 = 3 \tag{69}$$

Each eigenvector is the solution of eq.(64), which for μ_1 becomes:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{70}$$

This is two linearly dependent equations in two unknowns, and one solution is:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{71}$$

This is an (unnormalized) eigenvector corresponding to the first eigenvalue. The other eigenvector must be orthogonal to this, so it can be expressed as:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{72}$$

The eigenvectors in eqs.(71) and (72) can be normalized as follows:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad , \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{73}$$

So, the principal axes of the ellipse are directed at 45^{o} into the fourth quadrant (v_1) and first quadrant (v_2). The length of the 4th-quadrant semi-axis is $1/\sqrt{\mu_1} = 1$ and the length of the 1st-quadrant semi-axis is $1/\sqrt{\mu_2} = 1/\sqrt{3} \approx 0.58$. More explicitly, the semi-axes are:

$$s_1 = \frac{1}{\sqrt{\mu_1}} v_1 = \frac{1}{\sqrt{1}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad s_2 = \frac{1}{\sqrt{\mu_2}} v_2 = \frac{1}{\sqrt{3}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{74}$$