AMA1007 Supplementary Notes: To find the closest point from an ellipse to a given point

An ellipse in 2-D in standard position (centered at the origin and not tilted) can be expressed in the form of an implicit equation as

$$\frac{x^2}{r_1^2} + \frac{y^2}{r_2^2} = 1,$$

where r_1 and r_2 are the major and minor radius. In general, an ellipse not in standard position (and non-degenerate) can be expressed in the form of an implicit equation as

$$ax^2 + bxy + cy^2 + dx + ey + f = 1,$$

where $b^2 - 4ac < 0$.

An ellipse in 2-D not in standard position can also be expressed in parametric form:

$$\begin{aligned} x(t) &= r_1 \cos(\alpha)\cos(t) - r_2 \sin(\alpha)\sin(t) + c_x, \\ y(t) &= r_1 \sin(\alpha)\cos(t) + r_2 \cos(\alpha)\sin(t) + c_y, \end{aligned}$$

where t is the parameter varying in $[-\pi, \pi]$, the ellipse is centered at (c_x, c_y), and titled with an angle α with the x-axis measured in counter-clockwise direction. From the parametric form, the implicit equation can be recovered:

$$\begin{aligned} a &= r_1^2 \sin^2(\alpha) + r_2^2 \cos^2(\alpha) \\ b &= 2(r_2^2 - r_1^2)\sin(\alpha)\cos(\alpha) \\ c &= r_1^2 \cos^2(\alpha) + r_2^2 \sin^2(\alpha) \\ d &= -2ac_x - bc_y \\ e &= -bc_x - 2cc_y \\ f &= ac_x^2 + bc_xc_y + cc_y^2 - r_1^2 r_2^2 \end{aligned}$$

Consider a fixed point P_1 with coordinate (x_0, y_0). To find the closest point on the ellipse to P_1, we first formulate the square of the distance from a point on the ellipse to P_1 in parametric form

$$\begin{aligned} DS(t) &= (x(t) - x_0)^2 + (y(t) - y_0)^2 \\ &= (r_1 \cos(\alpha)\cos(t) - r_2 \sin(\alpha)\sin(t) + c_x - x_0)^2 \\ &\quad +(r_1 \sin(\alpha)\cos(t) + r_2 \cos(\alpha)\sin(t) + c_y - y_0)^2 \end{aligned}$$

Finding the candidates for global minimum of DS with respect to t, we solve

$$\frac{d}{dt}DS(t) = 0,$$

for $t = t_0, t_1, t_2, t_3$. Note, it is possible to have a total of

- 4 real roots within $-\pi \leq t \leq \pi$ (2 maxima (with which 1 is a global max and 1 is just a local max) and 2 minima (with which 1 is a global min and 1 is just a local min)), or
- 3 real roots (1 max, 1 min, and a double root neither max nor min), or
- 2 real roots (1 max and 1 min).

Then, we check to see if $\frac{d^2}{dt^2}DS(t_i) > 0$ for $i = 0, 1, 2, 3$, i.e, to see which are local minima. **Suppose two such local minima were found in the 4 real roots situation, we compare the value of DS to get the global minimum**.

When the solution (value of t_0 that minimize $DS(\cdot)$) cannot not be easily obtained analytically, one would then seek for a numerical solution for t_0.

Example Consider a fixed point P_1 given by $(3, 2)$, and an ellipse

$$\begin{aligned} x(t) &= 2\cos(\pi/4)\cos(t) - \sin(\pi/4)\sin(t) + 3, \\ y(t) &= 2\sin(\pi/4)\cos(t) + \cos(\pi/4)\sin(t). \end{aligned}$$

The square of the distance from a point on the ellipse to P_1 in parametric form

$$\begin{aligned} DS(t) &= (x(t) - 3)^2 + (y(t) - 2)^2 \\ &= (\sqrt{2}\cos(t) - \frac{\sqrt{2}}{2}\sin(t))^2 + (\sqrt{2}\cos(t) + \frac{\sqrt{2}}{2}\sin(t) - 2)^2. \end{aligned}$$

Solving $\frac{d}{dt}DS(t) = 0$ yields $\sin(t) = -\frac{2\cos(t)}{3\sqrt{2}\cos(t) - 4}$. We thus resolve to find the roots numerically.

Within $t \in (-\pi, \pi)$ there are two roots for $\frac{d}{dt}DS(t) = 0$, namely

$t_0 \approx 0.934295578$, and
$t_1 \approx -2.900080183$.

We further check that t_0 has positive second derivative, thus, t_0 is a minimum, whereas, t_1 has a negative second derivative, and thus, a maximum.

The closest point on the ellipse to P_1 is thus given by
$x(t_0) \approx 3.271925$ and $y(t_0) \approx 1.409229$.