The point-plane distance by two different routes

Introduction

Recently I was seeking a question to place on a final examination to test my students' mastery of a technique of optimization in the presence of constraints (Lagrange multipliers). I decided on the following question:

"Find the minimum value of $f(x, y, z) = (x-3)^2 + (y-4)^2 + (z+1)^2$ subject to the constraint $g(x, y, z) = 2x + y - 2z = 3$ **and** find the location (x, y, z) where this minimum value occurs."

This question finds the square of the distance between the point $P(3, 4, -1)$ and the plane $2x + y - 2z = 3$ and finds the closest point on the plane to *P*. It then reminded me of the very different method used in a previous course on linear algebra to solve similar questions, using the geometry of vectors.

That led me to tackle the general problem of the distance between any point *P* and any plane, together with the location of the point *Q* on the plane closest to the point *P*, by each of the distinct methods of vector geometry and calculus.

Method using vectors

A normal vector to the plane $\Pi: Ax + By + Cz + D = 0$ is

$$\vec{\mathbf{n}} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}, \quad \text{whose magnitude is} \quad n = |\vec{\mathbf{n}}| = \sqrt{A^2 + B^2 + C^2}$$

Let $Q(a, b, c)$ be the point on the plane closest to $P(x_0, y_0, z_0)$ and let $R(x, y, z)$ be a general point on the plane.

If *P* is on the plane, then the problem is trivial: the points *Q* and *P* are the same point and the distance from *P* to the plane is zero.

Otherwise, there are four possible distinct configurations:

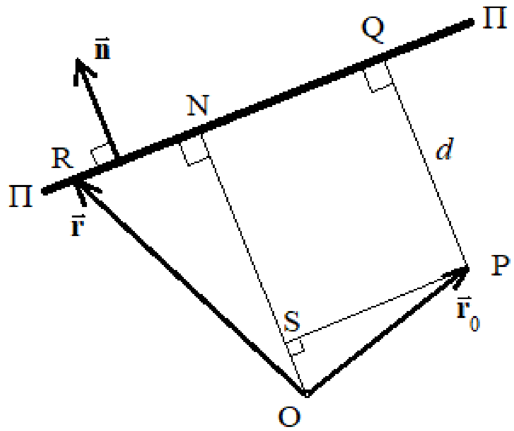


Figure 1

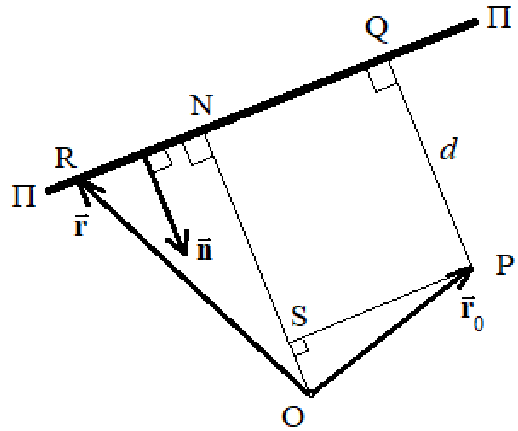


Figure 2

Figure 3

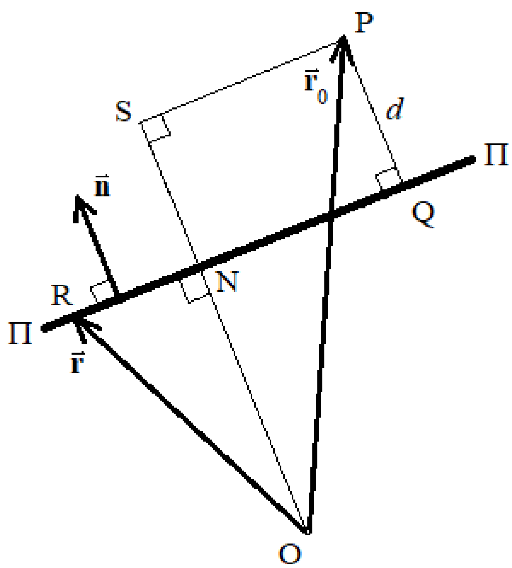
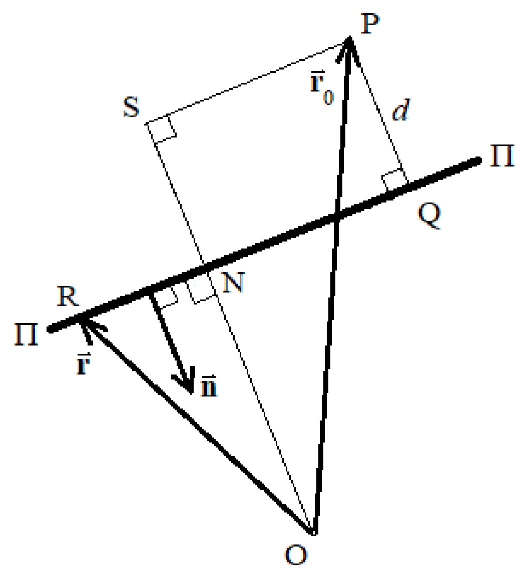


Figure 4



In all four configurations, $\vec{\mathbf{r}}\bullet\vec{\mathbf{n}} = Ax + By + Cz = -D$

In figures 1 and 3, ON is the projection of $\overrightarrow{OR} = \vec{\mathbf{r}}$ in the direction of $\vec{\mathbf{n}}$,

so that $ON = \vec{\mathbf{r}}\bullet\hat{\mathbf{n}} = \dfrac{\vec{\mathbf{r}}\bullet\vec{\mathbf{n}}}{n} = -\dfrac{D}{n}$

and the projection of $\overrightarrow{OP} = \vec{\mathbf{r}}_0$ in the direction of $\vec{\mathbf{n}}$ is

$$OS = \vec{\mathbf{r}}_0 \bullet \hat{\mathbf{n}} = \frac{\vec{\mathbf{r}}_0 \bullet \vec{\mathbf{n}}}{n} = \frac{Ax_0 + By_0 + Cz_0}{n}$$

In figures 2 and 4, the normal vector $\vec{\mathbf{n}}$ is pointing towards the origin relative to the plane and

$$ON = -\vec{\mathbf{r}} \bullet \hat{\mathbf{n}} = \frac{-\vec{\mathbf{r}} \bullet \vec{\mathbf{n}}}{n} = +\frac{D}{n} \quad \text{and} \quad OS = -\frac{Ax_0 + By_0 + Cz_0}{n}$$

In figures 1 and 2, $d = PQ = SN = ON - OS$

In figures 3 and 4, $d = QP = NS = OS - ON$

The results for d in these four scenarios are:

Figure 1: $d = -\dfrac{D}{n} - \dfrac{Ax_0 + By_0 + Cz_0}{n}$

Figure 2: $d = +\dfrac{D}{n} - \left(-\dfrac{Ax_0 + By_0 + Cz_0}{n} \right)$

Figure 3: $d = \dfrac{Ax_0 + By_0 + Cz_0}{n} - \left(-\dfrac{D}{n} \right)$

Figure 4: $d = -\dfrac{Ax_0 + By_0 + Cz_0}{n} - \dfrac{D}{n}$

In all four cases, this becomes $d = \dfrac{|\, Ax_0 + By_0 + Cz_0 + D \,|}{n} \quad \Rightarrow$

$$d = \frac{|\, Ax_0 + By_0 + Cz_0 + D \,|}{\sqrt{A^2 + B^2 + C^2}}$$

$\overrightarrow{OQ} = \overrightarrow{OP} + \overrightarrow{PQ} = \vec{\mathbf{r}}_0 \pm d\hat{\mathbf{n}}$, with the sign depending on the orientation of the normal vector and on whether the point *P* and the origin *O* are on the same side of the plane.

$$d\hat{\mathbf{n}} = \frac{|\, Ax_0 + By_0 + Cz_0 + D \,|}{\sqrt{A^2 + B^2 + C^2}} \cdot \frac{\vec{\mathbf{n}}}{n} = \frac{|\, Ax_0 + By_0 + Cz_0 + D \,|}{A^2 + B^2 + C^2} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

$Ax_0 + By_0 + Cz_0 + D > 0$ if either

$\vec{\mathbf{n}}$ is pointing towards the origin relative to the plane and *O*, *P* are on the same side of the plane (figure 2);

or

$\vec{\mathbf{n}}$ is pointing away from the origin relative to the plane and O, P are on opposite sides of the plane (figure 3).

Otherwise $Ax_0 + By_0 + Cz_0 + D \ < \ 0$ (figures 1 and 4)

In all four situations,

$$\overrightarrow{PQ} = -\frac{Ax_0 + By_0 + Cz_0 + D}{\sqrt{A^2 + B^2 + C^2}} \cdot \frac{\vec{\mathbf{n}}}{n} = -\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

A unique result for the location of the point Q follows:

$$\overrightarrow{OQ} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} - \left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2} \right) \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

Method using Lagrange multipliers

We are seeking the minimum value of

$f(x, y, z) = (x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2$

subject to the constraint $g(x, y, z) = Ax + By + Cz + D = 0$

Functions f and g are both polynomial functions of x, y and z and are therefore infinitely differentiable on all of $\mathbb{R}^2$. There are no boundary points to check. Obviously f can be made arbitrarily large by taking values of x, y and z that are large (in magnitude) such that $g = 0$. A unique extremum must therefore be an absolute minimum.

At any extremum $\vec{\nabla} f = \lambda \vec{\nabla} g \quad \Rightarrow \quad \begin{bmatrix} 2(x - x_o) \\ 2(y - y_o) \\ 2(z - z_o) \end{bmatrix} = \lambda \begin{bmatrix} A \\ B \\ C \end{bmatrix}$

Together with the constraint $g = 0$, this generates the system of simultaneous equations

$$2(x - x_o) = \lambda A \qquad (\mathbf{1})$$
$$2(y - y_o) = \lambda B \qquad (\mathbf{2})$$
$$2(z - z_o) = \lambda C \qquad (\mathbf{3})$$
$$Ax + By + Cz + D = 0 \qquad (\mathbf{4})$$

Provided $A \neq 0$, substitute **(1)** in each of **(2)** and **(3)**:

$$(y - y_o) = \frac{B}{A}(x - x_o) \quad \Rightarrow \quad y = y_o + \frac{B}{A}(x - x_o) \qquad (\mathbf{5})$$

and

$$(z - z_o) = \frac{C}{A}(x - x_o) \quad \Rightarrow \quad z = z_o + \frac{C}{A}(x - x_o) \qquad (\mathbf{6})$$

Substitute **(5)** and **(6)** in **(4)**:

$$Ax + B\left(y_o + \frac{B}{A}(x - x_o)\right) + C\left(z_o + \frac{C}{A}(x - x_o)\right) + D = 0$$
$$\Rightarrow \quad \left(A + \frac{B^2}{A} + \frac{C^2}{A}\right)x = -D - By_o - Cz_o + \left(\frac{B^2}{A} + \frac{C^2}{A}\right)x_o$$
$$\Rightarrow \quad \left(A^2 + B^2 + C^2\right)x = -AD - ABy_o - ACz_o + \left(B^2 + C^2\right)x_o$$
$$\Rightarrow \quad x = \frac{-AD - ABy_o - ACz_o + \left(A^2 + B^2 + C^2 - A^2\right)x_o}{A^2 + B^2 + C^2}$$
$$\Rightarrow \quad x = x_o - A\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)$$
$$(\mathbf{5}) \quad \Rightarrow \quad y = y_o + \frac{B}{A}(x - x_o) = y_o + \frac{B}{A}\left(-A\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)\right)$$
$$\Rightarrow \quad y = y_o - B\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)$$
$$(\mathbf{6}) \quad \Rightarrow \quad z = z_o + \frac{C}{A}(x - x_o) = z_o + \frac{C}{A}\left(-A\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)\right)$$
$$\Rightarrow \quad z = z_o - C\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)$$

Therefore the only critical point is at

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} - \left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2} \right) \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

which is the nearest point on the plane $Ax + By + Cz + D = 0$ to the point $P(x_o, y_o, z_o)$.

If $A = 0$, then one can modify this method to arrive at the same result, or one can argue from the symmetry in this expression that it is true even if any one or two of A, B, C are zero.

The minimum value of f is

$$f(x, y, z) = (x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2$$

$$= \left(-\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2} \right) \right)^2 \left(A^2 + B^2 + C^2 \right) \quad \Rightarrow$$

$$f_{\min} = d^2 = \frac{\left(Ax_o + By_o + Cz_o + D \right)^2}{A^2 + B^2 + C^2} \quad \Rightarrow$$

$$d = \frac{\left| Ax_o + By_o + Cz_o + D \right|}{\sqrt{A^2 + B^2 + C^2}}$$

Method by unconstrained optimization

Functions f and g are both polynomial functions of x and y and are therefore infinitely differentiable on all of $\mathbb{R}^2$. There are no boundary points to check.

Substitute $g(x, y, z) = Ax + By + Cz + D = 0$ into

$f(x, y, z) = (x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2$:

$z = \dfrac{-(Ax + By + D)}{C}$ (provided $C \neq 0$)

The function to be minimised is now

$$f(x,y) = (x-x_o)^2 + (y-y_o)^2 + \left(\frac{-(Ax+By+D)}{C} - z_o\right)^2$$

Critical points of $f(x,y)$ occur only where $\vec{\nabla} f = \vec{\mathbf{0}}$

$$\vec{\nabla} f = \begin{bmatrix} 2(x-x_o)+0+2\left(\frac{-(Ax+By+D)}{C}-z_o\right)\left(\frac{-(A+0+0)}{C}-0\right) \\ 0+2(y-y_o)+2\left(\frac{-(Ax+By+D)}{C}-z_o\right)\left(\frac{-(0+B+0)}{C}-0\right) \end{bmatrix}$$

$$= \frac{2}{C^2}\begin{bmatrix} C^2(x-x_o) - A\left(-(Ax+By+D)-Cz_o\right) \\ C^2(y-y_o) - B\left(-(Ax+By+D)-Cz_o\right) \end{bmatrix}$$

$$= \frac{2}{C^2}\begin{bmatrix} (A^2+C^2)x + ABy + C(Az_o - Cx_o) + AD \\ ABx + (B^2+C^2)y + C(Bz_o - Cy_o) + BD \end{bmatrix}$$

$$\vec{\nabla} f = \vec{\mathbf{0}} \quad \Rightarrow \quad M \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -AD + C(Cx_o - Az_o) \\ -BD + C(Cy_o - Bz_o) \end{bmatrix}$$

where $M = \begin{bmatrix} A^2+C^2 & AB \\ AB & B^2+C^2 \end{bmatrix}$

$$\det M = (A^2+C^2)(B^2+C^2) - (AB)^2 = C^2(A^2+B^2+C^2)$$

$$M^{-1} = \frac{1}{C^2(A^2+B^2+C^2)}\begin{bmatrix} B^2+C^2 & -AB \\ -AB & A^2+C^2 \end{bmatrix}$$

$$\Rightarrow \quad \begin{bmatrix} x \\ y \end{bmatrix} = M^{-1}\begin{bmatrix} -AD + C(Cx_o - Az_o) \\ -BD + C(Cy_o - Bz_o) \end{bmatrix}$$

$$= \frac{1}{C^2(A^2+B^2+C^2)} \times$$

$$\times \begin{bmatrix} (B^2+C^2)\left(-AD + C(Cx_o - Az_o)\right) - AB\left(-BD + C(Cy_o - Bz_o)\right) \\ -AB\left(-AD + C(Cx_o - Az_o)\right) + (A^2+C^2)\left(-BD + C(Cy_o - Bz_o)\right) \end{bmatrix}$$

$$= \frac{1}{C^2\left(A^2+B^2+C^2\right)}\begin{bmatrix} -C^2AD + C^2\left(\left(B^2+C^2\right)x_o - ABy_o - ACz_o\right) \\ -C^2BD + C^2\left(\left(A^2+C^2\right)y_o - ABx_o - BCz_o\right) \end{bmatrix}$$

$$= \frac{1}{A^2+B^2+C^2}\begin{bmatrix} -A\left(D+By_o+Cz_o\right)+\left(A^2+B^2+C^2-A^2\right)x_o \\ -B\left(D+Ax_o+Cz_o\right)+\left(A^2+B^2+C^2-B^2\right)y_o \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \end{bmatrix} + \frac{1}{A^2+B^2+C^2}\begin{bmatrix} -A\left(Ax_o+By_o+Cz_o+D\right) \\ -B\left(Ax_o+By_o+Cz_o+D\right) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \end{bmatrix} - \frac{Ax_o+By_o+Cz_o+D}{A^2+B^2+C^2}\begin{bmatrix} A \\ B \end{bmatrix}$$

and $z = \dfrac{-\left(Ax+By+D\right)}{C}$

$$= \frac{-\left(Ax_o+By_o+D\right)}{C} + \frac{Ax_o+By_o+Cz_o+D}{C\left(A^2+B^2+C^2\right)}\left(A^2+B^2\right)$$

$$= \frac{-\left(Ax_o+By_o+D\right)}{C} + \frac{Ax_o+By_o+Cz_o+D}{C\left(A^2+B^2+C^2\right)}\left(A^2+B^2+C^2-C^2\right)$$

$$= \frac{-\left(Ax_o+By_o+D\right)}{C} + \frac{Ax_o+By_o+Cz_o+D}{C} - \frac{C^2\left(Ax_o+By_o+Cz_o+D\right)}{C\left(A^2+B^2+C^2\right)}$$

$$= z_o - \frac{C\left(Ax_o+By_o+Cz_o+D\right)}{A^2+B^2+C^2}$$

Therefore the only critical point is at

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} - \left(\frac{Ax_o+By_o+Cz_o+D}{A^2+B^2+C^2}\right)\begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

which is the nearest point on the plane $Ax+By+Cz+D=0$ to the point $P\left(x_o, y_o, z_o\right)$.

In the event that $C = 0$, then express y in terms of x and substitute $y = -\dfrac{Ax + D}{B}$ into $f(x, y)$ in the analysis above, unless B and C are both zero, in which case substitute $x = -\dfrac{D}{A}$ into $f(x, y)$. Alternatively, one can make an argument based on symmetry that this expression for the location of the critical point is valid even if any one or two of A, B, C are zero.

As before, the minimum value of f is

$$f(x, y, z) = (x - x_o)^2 + (y - y_o)^2 + (z - z_o)^2$$

$$= \left(-\left(\frac{Ax_o + By_o + Cz_o + D}{A^2 + B^2 + C^2}\right)\right)^2 \left(A^2 + B^2 + C^2\right) \quad \Rightarrow$$

$$f_{\min} = d^2 = \frac{\left(Ax_o + By_o + Cz_o + D\right)^2}{A^2 + B^2 + C^2} \quad \Rightarrow$$

$$d = \frac{\left| Ax_o + By_o + Cz_o + D \right|}{\sqrt{A^2 + B^2 + C^2}}$$

This method requires more effort than the method using Lagrange multipliers, while the vector approach is the fastest (and easiest to visualize) of these three methods.

These general results are incorporated easily into a spreadsheet, which helps in the setting of sample exercises and examination problems. From an Excel spreadsheet [1], the solution to the examination question at the top of this note is quickly found to be $f_{\min} = 9 \Rightarrow d = 3$, with the nearest point at (1, 3, 1).

Reference

1. Available at http://www.engr.mun.ca/~ggeorge/point-plane.xlsx

Glyn George
Department of Electrical and Computer Engineering
Memorial University of Newfoundland
St. John's NL
Canada A1B 3X5

e-mail: glyn@mun.ca
web site: http://www.engr.mun.ca/~ggeorge