Point and line

We are asked to find an equation for the closest distance from a line to a point. Let the line have equation $Ax + By + C = 0$ and the point be (x_0, y_0).

I find it more familiar to work with $y = mx + b$ so we do

$$y = -\frac{A}{B}x - \frac{C}{B}$$

and then $m = -A/B$ and $b = -C/B$.

The first line has slope m, so llines perpendicular to it have slope $-1/m$. The particular one which goes through the point (x_0, y_0) is

$$y = -\frac{1}{m}x + \frac{x_0}{m} + y_0$$

The point at the intersection has

$$\begin{aligned} mx + b = y &= -\frac{1}{m}x + \frac{x_0}{m} + y_0 \\ m^2x + mb &= -x + x_0 + my_0 \\ (m^2 + 1)x &= x_0 + m(y_0 - b) \\ x &= \frac{x_0 + m(y_0 - b)}{m^2 + 1} \end{aligned}$$

The corresponding y is

$$y = \frac{mx_0 + m^2(y_0 - b)}{m^2 + 1} + b$$

The square of the distance is

$$d^2 = (x - x_0)^2 + (y - y_0)^2$$

$$= \left[\ \frac{x_0 + m(y_0 - b)}{m^2 + 1} - x_0\ \right]^2 + \left[\frac{mx_0 + m^2(y_0 - b)}{m^2 + 1} + b - y_0\ \right]^2$$

which looks a mess. However, when we put things over the common denominator we get some cancelations:

$$d^2 = \left[\ \frac{m(y_0 - b) - x_0 m^2)}{m^2 + 1}\ \right]^2 + \left[\frac{mx_0 + b - y_0}{m^2 + 1}\ \right]^2$$

Factoring out $-m$:

$$= \left[\ (-m) \cdot \frac{(mx_0 - y_0 + b)}{m^2 + 1}\ \right]^2 + \left[\frac{mx_0 - y_0 + b}{m^2 + 1}\ \right]^2$$

and then the wonderful cancelation since we also have $m^2 + 1$ on top

$$= \frac{(mx_0 - y_0 + b)^2}{m^2 + 1}$$

Taking the negative square root for a reason we'll see momentarily

$$d = \frac{-mx_0 + y_0 - b}{\sqrt{1 + m^2}}$$

Substituting for $m = -A/B$ and $b = -C/B$:

$$d = \frac{\frac{A}{B}x_0 + y_0 + \frac{C}{B}}{\sqrt{1 + \frac{A^2}{B^2}}}$$

Multiply top and bottom by B:

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

In other words, substitute the point (x_0, y_0) into the original equation of the line and divide by $\sqrt{A^2 + B^2}$. Pretty simple.

We have ignored the complication that one line or the other may be vertical, in which case this derivation isn't valid. Kline says the equation is still good even in that case.

by rotation

Another approach is to rotate the system. We are given the line with equation

$$Ax + By + C = 0$$

The slope of the line is $m = -A/B$. We need to rotate by an angle θ whose tangent is equal to that.

$$\theta = \tan^{-1} -A/B$$

The rotation equations are:

$$x' = x\cos\theta - y\sin\theta$$

$$y' = x\sin\theta + y\cos\theta$$

So the rotated line is

$$A(x\cos\theta - y\sin\theta) + B(x\sin\theta + y\cos\theta) + C = 0$$

We can solve for y:

$$y = \frac{(A\cos\theta + B\sin\theta)x + C}{A\sin\theta - B\cos\theta}$$

We must also rotate the point (x_0, y_0):

$$x_0' = x_0\cos\theta - y_0\sin\theta$$

$$y_0' = x_0 \sin\theta + y_0 \cos\theta$$

We find the point on the line corresponding to $x = x_0'$:

$$y = \frac{(A\cos\theta + B\sin\theta)(x_0\cos\theta - y_0\sin\theta) + C}{A\sin\theta - B\cos\theta}$$

and simply subtract y_0':

$$d = \frac{(A\cos\theta + B\sin\theta)(x_0\cos\theta - y_0\sin\theta) + C}{A\sin\theta - B\cos\theta} - x_0\sin\theta - y_0\cos\theta$$

Not sure this is correct but suppose

$$\theta = \tan^{-1} -A/B$$

then we would have

$$\sin\theta = A \qquad \cos\theta = -B$$

so we have on the bottom

$$d = \frac{(A\cos\theta + B\sin\theta)(x_0\cos\theta - y_0\sin\theta) + C}{A^2 + B^2} - x_0\sin\theta - y_0\cos\theta$$

which matches our previous result.

For the moment, leave it as it was

$$d = \frac{(A\cos\theta + B\sin\theta)(x_0\cos\theta - y_0\sin\theta) + C}{A\sin\theta - B\cos\theta} - x_0\sin\theta - y_0\cos\theta$$

The first part of the right-hand side is

$$Ax_0\cos^2\theta - Ay_0\sin\theta\cos\theta + Bx_0\sin\theta\cos\theta - By_0\sin^2\theta + C$$

and the rest is

$$-Ax_0\sin^2\theta - Ay_0\sin\theta\cos\theta + Bx_0\sin\theta\cos\theta + By_0\cos^2\theta$$

Adding, we have

$$\frac{(Ax_0 + By_0)\cos 2\theta + (Bx_0 - Ay_0)\sin 2\theta + C}{A^2 + B^2}$$

So even if we could find the error, it is definitely not any simpler.