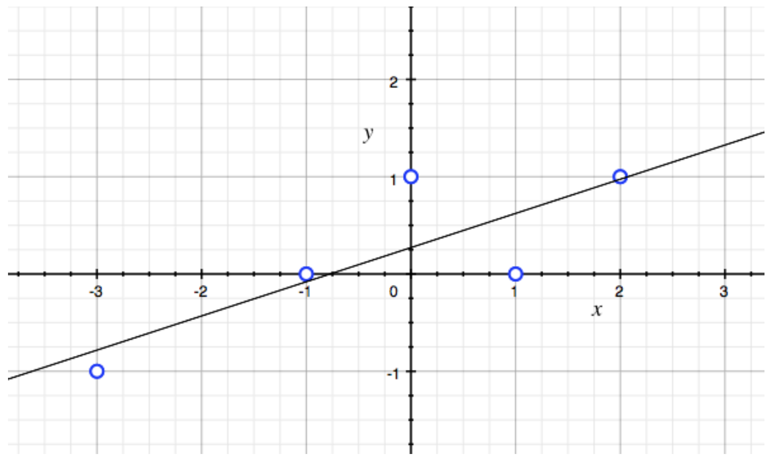
14.7-8: Optimization

Friday, March 11

Return of the Best Fit Line

Given the five points (-3,-1), (0,1), (-1,0), (1,0), and (2,1), we would like to find the line of the form $y = ax+b$ that best approximates the data. If our goal is to minimize the sum of the squared errors, our error function will be $E(a,b) = \sum_i (ax_i + b - y_i)^2 = 15a^2 + 5b^2 + 3 - 2ab - 2b - 10a$. Find the optimal pair (a,b) and plot the appropriate line.



$E_a = 30a - 2b - 10$ and $E_b = 10b - 2a - 2$, so setting $\nabla E(a,b) = 0$ gives the solution $(a,b) = (13/37, 10/37)$. The line looks pretty good.

Maxima and Minima

Find the shortest distance from the point $(2,0,-3)$ to the plane $x+y+z=1$.

If we make the substitution $z = 1-x-y$ then this constrained problem in 3 variables becomes an unconstrained problem in 2 variables: minimize $\sqrt{(2-x)^2+(-y)^2+(-3-(1-x-y))^2}$. Minimizing this function will have the same solution (x,y) as minimizing its square, which is $(2-x)^2+y^2+(x+y-4)^2$, or $2x^2+2y^2+20-12x-8y+2xy$.

Setting the gradient to zero gives the equations $4x-12+2y=0, 4y-8+2x=0$, or $2x+y=6, 2y+x=4$, with solution $(x,y) = (8/3, 2/3)$. Therefore $z = -7/3$.

This makes the shortest distance $\sqrt{(2/3)^2+(2/3)^2+(2/3)^2} = 2/\sqrt{3}$.

Minimize the function $f(x,y) = x^2+3y^2-4x-12y+16$ given the constraints $-1 \le x \le 1, -1 \le y \le 1$.

The gradient is $\langle 2x-4, 6y-12 \rangle$ and so the only critical point is $(2,2)$. This is outside of the constraint set, so we have to check the borders for the minimum. At $y=1, g(x) := f(x,1) = x^2-4x+7$ and has derivative $2x-4$. The minimum is at $x=2$, but this is again outside the contraint set. Similarly, the minimum does not occur along any of the other edges, so it must be at one of the corners. Checking manually shows that the minimum occurs at $(1,1)$.

Lagrange Multipliers

Find the maximum and minimum attainable values of $f(x,y) = xy$ subject to the contraint $4x^2 + y^2 = 8$.
$\nabla f = \lambda y, x\rangle$ and if $g(x,y) = 4x^2 + y^2$ then $\nabla g = \langle 8x, 2y\rangle = 2\langle 4x, y\rangle$, so using Lagrange multipliers gives the three constraints

$$\begin{aligned} y &= 4\lambda x \\ x &= \lambda y \\ 4x^2 + y^2 &= 8. \end{aligned}$$

Making the substitution $y = 4\lambda x$ in the second equation gives $x = 4\lambda^2 x$, so $\lambda = \pm 1/2$. In either case, $x = \pm y/2$, so $8 = 4x^2 + y^2 = 4(y/2)^2 + y^2 = 2y^2$, meaning $y = \pm 2$ and $x = \pm 1$. The four critical points are at $(\pm 1, \pm 2)$, and checking the values at these points shows that $(1,2)$ and $(-1,-2)$ are local maxes while $(-1,2)$ and $(1,-2)$ are local minima.

Find the points on the ellipse $(x-1)^2 + 4(y-2)^2 = 1$ with the maximum and minimum distances from the origin.
Look at the squared distance: if $f(x,y) = \sqrt{x^2+y^2}$, then $\nabla f^2(x,y) = 2\langle x, y\rangle$. Then if $g(x,y) = (x-1)^2 + 4(y-2)^2$, $\nabla g(x,y) = 2\langle x-1, 4y-8\rangle$. Using Lagrange multipliers then gives the three equations

$$\begin{aligned} x &= \lambda(x-1) \\ y &= \lambda(4y-8) \\ (x-1)^2 + 4(y-2)^2 &= 1. \end{aligned}$$

Then this gets icky. Solving for x in the first equation gives $x - 1 = \frac{1}{\lambda - 1}$ and solving for y in the second gives $y - 2 = 2/(4\lambda - 1)$, so substituting both into the third equation gives

$$1/(\lambda-1)^2 + 16/(4\lambda-1)^2 = 1,$$

which has solutions $\lambda = 5/8 \pm 3\sqrt{17}/8$, leading to the solutions $(x,y) = (\frac{7}{6} \mp \frac{\sqrt{17}}{6}, \frac{23}{12} \mp \frac{\sqrt{17}12}{)}$. ...okay, that was a little hard to work out by hand.

Find the shortest distance from the point $(2,0,-3)$ to the plane $x+y+z=1$, this time by using Lagrange multipliers. What does this have to do with the distance formulas from chapter 12?
Work with the squared distance, since the problems are equivalent. If $f(x,y,z) = (x-2)^2 + y^2 + (z+3)^2$ then $\nabla f = 2\langle x-2, y, z+3\rangle$. If $g(x,y,z) = x+y+z$ then $\nabla g = \langle 1,1,1\rangle$.
Then using Lagrange multipliers gives the set of equations

$$\begin{aligned} x - 2 &= \lambda \\ y &= \lambda \\ z + 3 &= \lambda \\ x + y + z &= 1, \end{aligned}$$

and adding the first three equations together gives $x + y + z = 3\lambda - 1$, so $\lambda = 2/3$. Therefore $(x,y,z) = (-8/3, 2/3, -7/3)$.
$\langle 1,1,1\rangle$ is the normal vector to the plane and $\nabla f(x)$ is parallel to the vector $|x - (2,0,-3)|$, so in geometric terms this is saying that the closest point is the one that differs from $(2,0,-3)$ only by some multiple of the normal vector to the plane, which is what we concluded in chapter 12.