Click here to visit folder of solution videos on course Canvas/Panopto page

Problem 1: A best fit line

The collection of 5 data points $(-1, 6)$, $(0, 3)$, $(1, 0)$, $(2, -3)$, $(3, -4)$ lies close to a line of negative slope; see Figure 1. We are going to compute that line.

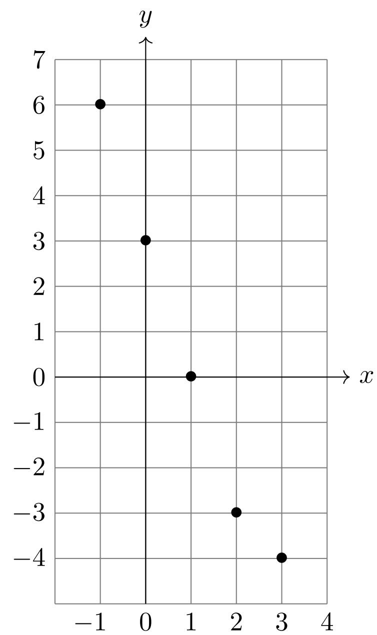


Figure 1: Five data points: $(-1, 6)$, $(0, 3)$, $(1, 0)$, $(2, -3)$, $(3, -4)$.

Suppose the line of best fit (in the least squares sense) is written as $y = mx + b$.

(a) Write down explicit 5-vectors $\mathbf{X}$ and $\mathbf{Y}$ so that for the 5-vector $\mathbf{1}$ whose entries are all equal to 1, the projection of $\mathbf{Y}$ into the plane $V = \mathrm{span}(\mathbf{X}, \mathbf{1})$ in $\mathbf{R}^5$ is $m\mathbf{X} + b\mathbf{1}$.

(b) Compute an orthogonal basis of $V = \mathrm{span}(\mathbf{X}, \mathbf{1})$ having the form $\{\mathbf{1}, \mathbf{v}\}$ for a 5-vector $\mathbf{v}$, and find scalars t and s so that $\mathbf{Proj}_V(\mathbf{Y}) = t\mathbf{v} + s\mathbf{1}$.

(c) By expressing $\mathbf{v}$ from (b) as a linear combination of $\mathbf{X}$ and $\mathbf{1}$, use your answer to (b) to find m and b so that the equation $y = mx + b$ gives the line of best fit. (As a safety check on your computations, you may want to plot your line on the above figure to see that it is a good fit for the data.)

Solution:

(a) The vectors are $\mathbf{X} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} 6 \\ 3 \\ 0 \\ -3 \\ -4 \end{bmatrix}$. (Given such vectors, the "m" and "b" given in the equation of the line of best fit are precisely those values for which the magnitude of the "vector of errors," i.e., $\|\mathbf{Y} - (m\mathbf{X} + b\mathbf{1})\|$, is minimized. This is in turn just another way of saying that $m\mathbf{X} + b\mathbf{1}$ is the vector in $V = \mathrm{span}(\mathbf{X}, \mathbf{1})$ which is closest to $\mathbf{Y}$; by definition, this is the projection of $\mathbf{Y}$ into V.)

(Note: entries can be rearranged, provided that both $\mathbf{X}$ and $\mathbf{Y}$ are rearranged in the same way.)

(b) The vector $\mathbf{v}$ can be taken to be $\mathbf{X} - \mathbf{Proj_1}(\mathbf{X})$, by the properties of the orthogonal projection. Furthermore,

$$\mathbf{Proj_1}(\mathbf{X}) = \left(\frac{\mathbf{X}\cdot\mathbf{1}}{\mathbf{1}\cdot\mathbf{1}}\right)\mathbf{1} = \overline{x}\mathbf{1},$$

with $\overline{x}$ equal to the average of the entries x_i in $\mathbf{X}$. (This is true because the numerator $\mathbf{X}\cdot\mathbf{1}$ is the sum of the entries of $\mathbf{X}$, while the denominator $\mathbf{1}\cdot\mathbf{1}$ is the number of entries, or 5.) We have the average $\overline{x} = (1/5)(-1+0+1+2+3) = 1$, so $\mathbf{v} = \mathbf{X} - (1)\mathbf{1} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$. (This $\mathbf{v}$ is what is called $\widehat{\mathbf{X}}$ in the course text.)

Now that we have an orthogonal basis $\{\mathbf{1}, \mathbf{v}\}$ for V, we may compute $\mathbf{Proj}_V(\mathbf{Y})$ (using the formula given in the Orthogonal Projection Theorem):

$$\mathbf{Proj}_V(\mathbf{Y}) = \mathbf{Proj_v}(\mathbf{Y}) + \mathbf{Proj_1}(\mathbf{Y}) = \left(\frac{\mathbf{Y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v} + \left(\frac{\mathbf{Y}\cdot\mathbf{1}}{\mathbf{1}\cdot\mathbf{1}}\right)\mathbf{1} = \left(\frac{\mathbf{Y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v} + \overline{y}\mathbf{1}.$$

We compute the dot products $\mathbf{Y}\cdot\mathbf{v} = -12 - 3 - 3 - 8 = -26$ and $\mathbf{v}\cdot\mathbf{v} = 4 + 1 + 0 + 1 + 4 = 10$; and also $\overline{y} = 2/5 = 0.4$. Thus, $\mathbf{Proj}_V(\mathbf{Y}) = -(2.6)\mathbf{v} + (0.4)\mathbf{1}$. (That is, $t = -2.6$ and $s = 0.4$.)

(c) Using the fact from (b) that $\mathbf{v} = \mathbf{X} - \mathbf{1}$, we may rewrite our computed linear combination for $\mathbf{Proj}_V(\mathbf{Y})$ as a combination of $\mathbf{X}$ and $\mathbf{1}$, with the final outcome being $m\mathbf{X} + b\mathbf{1}$ for the sought-after coefficients m and b:

$$\begin{aligned} \mathbf{Proj}_V(\mathbf{Y}) &= -(2.6)\mathbf{v} + (0.4)\mathbf{1} \\ &= -(2.6)(\mathbf{X} - \mathbf{1}) + (0.4)\mathbf{1} \\ &= -(2.6)\mathbf{X} + (2.6)\mathbf{1} + (0.4)\mathbf{1} \\ &= -(2.6)\mathbf{X} + (2.6 + 0.4)\mathbf{1} = -(2.6)\mathbf{X} + (3)\mathbf{1}. \end{aligned}$$

Hence, the line of best fit is $y = -(2.6)x + 3$. (That is, $m = -2.6$ and $b = 3$.) A plot of the data shows it is reasonably close to that line.

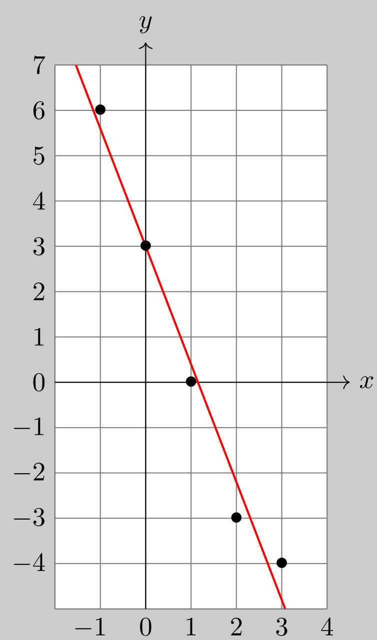


Figure 2: Line of best fit for given data.

Problem 2: Level sets of multivariable functions

(a) Describe and sketch the level sets of $\ln(y - x^2)$ on the region where $y > x^2$, relating each level set to the parabola $y = x^2$.

(b) Describe and sketch the level sets of $\cos(x^2 + y^2)$.

(c) Express the surface graph of $f(x, y) = x^2 + y^2$ in $\mathbf{R}^3$ as a level set of a function $h(x, y, z)$.

(d) (Extra) By using polar coordinates, describe the part of the graph of $f(x, y) = x^2 + y^2$ from (c) that lies over a line in the xy-plane through the origin, and use that to sketch the actual surface graph. (Don't "cheat" by looking on a computer; the point is to learn for yourself how to use restriction over well-chosen lower-dimensional subspaces, such as lines through the origin in $\mathbf{R}^2$, to build up a mental model of what happens over the entire domain.)

Solution:

(a) The region $y > x^2$ is the part of $\mathbf{R}^2$ "above" the parabola $y = x^2$ (away from the shaded region in Figure 3). The level sets correspond to the condition $\ln(y - x^2) = c$, or in other words $y - x^2 = e^c$, which is to say $y = x^2 + e^c$ with any $c \in \mathbf{R}$. The number e^c can be any positive value, so the level sets are the parabolas $y = x^2 + a$ for $a > 0$; this is the collection of "nested" parabolas obtained by translating $y = x^2$ upwards; some of these are shown (for various $c = \ln(a)$) in different colors in Figure 3.

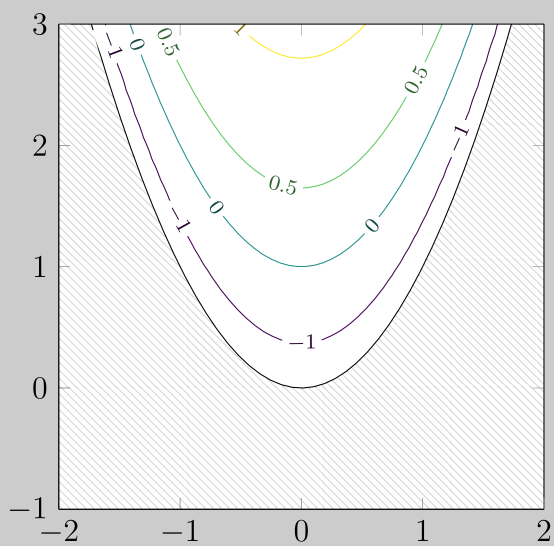


Figure 3: The level sets $\ln(y - x^2) = c$ for various c.

(b) The level sets are $\cos(x^2 + y^2) = c$ for $-1 \leq c \leq 1$. For each c, the corresponding level set is an infinite collection of circles centered at the origin (including the origin as a point when $c = 1$), with $x^2 + y^2$ varying through the values which $\cos$ carries to c (and the radii of those circles being the square roots of the values for $x^2 + y^2$ on the level set).

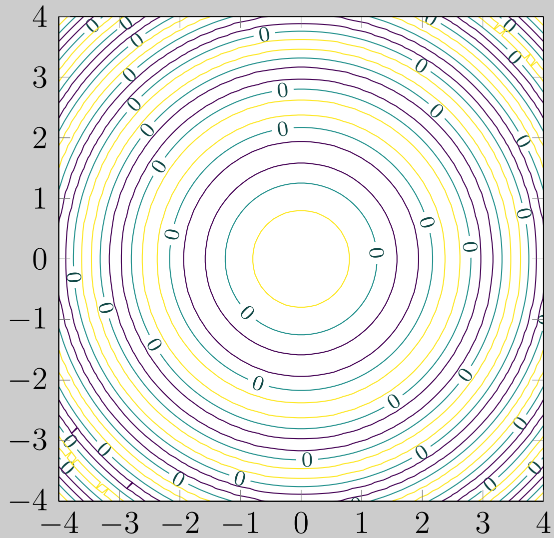


Figure 4: The level sets of $\cos(x^2+y^2)$ are nested circles (those of a fixed color).

(c) The graph is the region $z = x^2 + y^2$ in $\mathbf{R}^3$, so this is the level set $z - (x^2 + y^2) = 0$ for the function $h(x, y, z) = z - (x^2 + y^2)$.

(d) The part of the graph over a line in the xy-plane making an angle θ with the x-axis is given by substituting into $f(x, y)$ the expressions $x = \pm r\cos\theta$ and $y = \pm r\sin\theta$ (using a common sign: we need to allow the sign to account for both halves of the line, on either side of the origin). This yields $(\pm r\cos\theta)^2 + (\pm r\sin\theta)^2 = r^2$ over the two points at a distance r from the origin along that line in both directions, so it is the graph of the parabola $g(r) = r^2$ as we let $r \geq 0$ vary.

This parabola is the outcome regardless of the angle θ! Hence, it says that the surface graph is obtained by spinning that parabola in the xz-plane ($\theta = 0$) around the z-axis (to sweep out all possible angles θ), so it is a "cone-like" shape but with round bottom at the origin (rather than a sharp tip). Figure 5 shows the surface graph and the parabolic slices of the surface graph (in red, green, and blue) for several angles θ. If you zoom in, you'll see that these parabolas all meet at a common point, which is the bottom of the rounded tip. (It may look as if the red and green parabolas have their minimum points somewhere else, but that is an optical illusion caused by a rendering a 3-dimensional situation on a flat piece of paper or computer screen.)

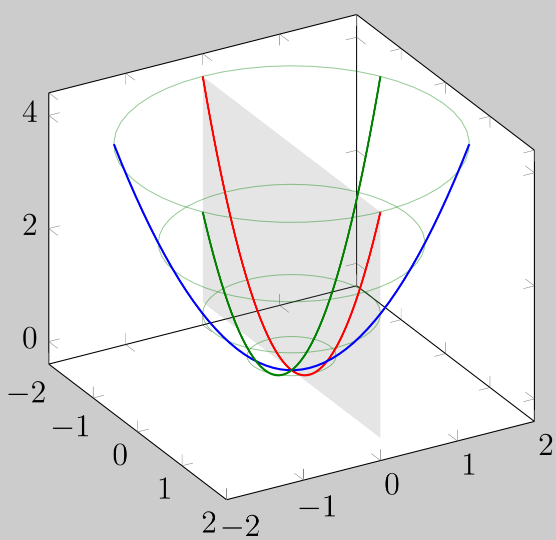


Figure 5: The part of $z = x^2 + y^2$ lying over a line (through the origin) in the xy plane is always a parabola.

Problem 3: Computations with vector-valued functions

For the functions $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ and $\mathbf{g} : \mathbf{R}^m \to \mathbf{R}^p$ below, compute $\mathbf{g} \circ \mathbf{f} : \mathbf{R}^n \to \mathbf{R}^p$ by working out its component functions; in each part also state the values of n, m, and p.

(a) $\mathbf{f}(x, y) = (e^x \cos(y), e^x \sin(y))$, $\mathbf{g}(v, w) = (v^2 - w^2, 2vw)$

(b) $\mathbf{f}(x, y) = (x^2 - y^2, 2xy)$, $\mathbf{g}(v, w) = (e^v \cos(w), e^v \sin(w))$

(c) $\mathbf{f}(t) = (1 - t^2, 2t, 1 + t^2)$, $\mathbf{g}(x, y, z) = x^2 + y^2 - z^2$

Solution:

(a) Here $\mathbf{f} : \mathbf{R}^2 \to \mathbf{R}^2$, so $n = m = 2$; and $\mathbf{g} : \mathbf{R}^2 \to \mathbf{R}^2$, so $p = 2$. For $\mathbf{g} \circ \mathbf{f}$, we have

$$\begin{aligned} \mathbf{g}(\mathbf{f}(x, y)) = \mathbf{g}(e^x \cos(y), e^x \sin(y)) &= ((e^x \cos(y))^2 - (e^x \sin(y))^2, 2e^x \cos(y)e^x \sin(y)) \\ &= (e^{2x}(\cos^2(y) - \sin^2(y)), e^{2x}(2\cos(y)\sin(y))) \\ &= (e^{2x} \cos(2y), e^{2x} \sin(2y)). \end{aligned}$$

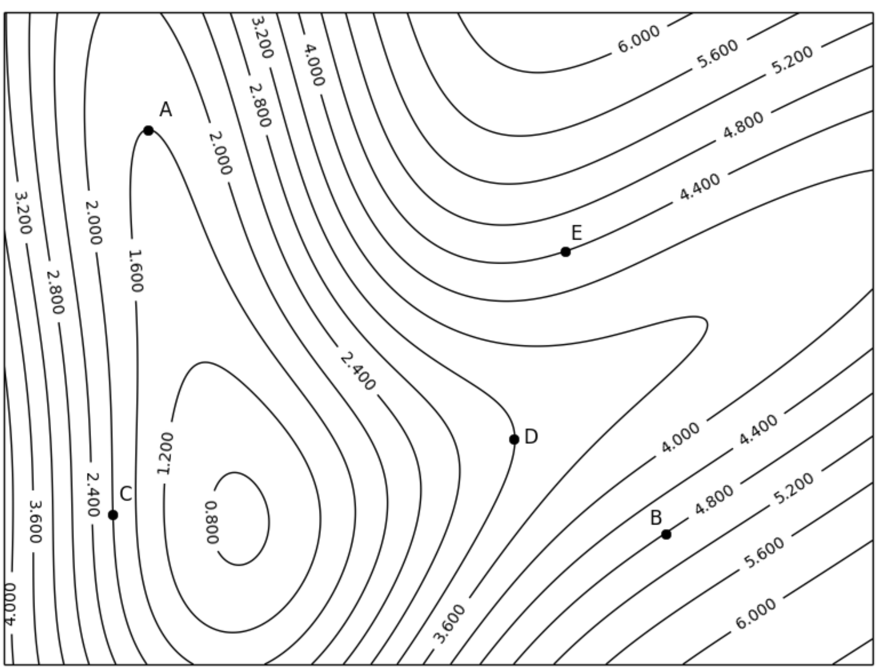
(It is unnecessary to make the final observation via double-angle formulas, but is geometrically informative.)

(b) Here the roles of $\mathbf{f}$, $\mathbf{g}$ are simply interchanged from those in (a); so again $n = m = p = 2$. However, for $\mathbf{g} \circ \mathbf{f}$ we now have $\mathbf{g}(\mathbf{f}(x, y)) = \mathbf{g}(x^2 - y^2, 2xy) = (e^{x^2-y^2} \cos(2xy), e^{x^2-y^2} \sin(2xy))$. This is very different from (a)!

(c) Here $\mathbf{f} : \mathbf{R} \to \mathbf{R}^3$, so $n = 1$ and $m = 3$; meanwhile $\mathbf{g} : \mathbf{R}^3 \to \mathbf{R}$, so $p = 1$. For $\mathbf{g} \circ \mathbf{f}$ (a scalar-valued function of one variable), we have $\mathbf{g}(\mathbf{f}(t)) = \mathbf{g}(1-t^2, 2t, 1+t^2) = (1-t^2)^2+(2t)^2-(1+t^2)^2 = 1-2t^2+t^4+4t^2-(1+2t^2+t^4) = 0$.

Problem 4: Visually interpreting derivatives

Below is a collection of level sets of a function $f : \mathbf{R}^2 \to \mathbf{R}$. (As usual, x is horizontal and y is vertical, and the length scales in the x- and y-directions are equal.)



(a) (Choose one) $\dfrac{\partial f}{\partial y}$ at **A** is: NEGATIVE ZERO $\boxed{\text{POSITIVE}}$

(b) (Choose one) $\dfrac{\partial f}{\partial y}$ at **B** is: $\boxed{\text{NEGATIVE}}$ ZERO POSITIVE

(c) (Choose one) $\dfrac{\partial f}{\partial y}$ at **C** is: NEGATIVE $\boxed{\text{ZERO}}$ POSITIVE

(d) (Choose one) $\dfrac{\partial f}{\partial x}$ at **D** is: NEGATIVE ZERO $\boxed{\text{POSITIVE}}$

(e) Which partial derivative is larger, in *absolute value*? $|f_y(\mathbf{A})|$ $\boxed{|f_y(\mathbf{B})|}$

(f) Which partial derivative is larger, in *absolute value*? $|f_x(\mathbf{E})|$ $\boxed{|f_y(\mathbf{E})|}$

(g) At what point(s) (not necessarily labeled) in the region depicted does f reasonably seem to have a local minimum? a local maximum? What can you say about the value taken by f at each of these points?

A picture to help you visualize contour plots. This is not part of the problem.

Solution:

(a) Answer: positive.

Moving a small distance in the positive y ("north") direction from **A** changes the value of f from 1.6 (precisely) to some value in the range between 1.6 and 2.0; thus, f increases.

(b) Answer: negative.

Moving a small distance in the positive y ("north") direction from **B** changes the value of f from 4.8 (precisely) to some value in the range between 4.4 and 4.8; thus, f decreases.

(c) Answer: zero.

Moving a small distance in the positive y ("north") direction from **C** is simply moving along (i.e., tangent to) the level set of f at level 2.0; thus, f is (instantaneously) unchanging.

(d) Answer: positive.

Moving a small distance in the positive x ("east") direction from **D** changes the value of f from 3.2 (precisely) to some value in the range between 3.2 and 3.6; thus, f increases.

(e) Answer: $|f_y(\mathbf{B})|$.

Consecutive level curves are vertically spaced much closer together near **B** than near **A**; so the (absolute value of the) vertical rate of change at **B** is greater than at **A**.

(f) Answer: $|f_y(\mathbf{E})|$.

The *vertical* spacing of consecutive level curves near **E** is closer together than the *horizontal* spacing here; so the (absolute value of the) vertical rate of change is greater than the (absolute value of the) horizontal rate of change.

(g) We expect that f has a local extremum at some point *inside* the loop, labeled "0.800," that is located to the right of point **C**. Judging from the decreasing label-values near this point, we conclude this extremum is a local minimum for f, with f-value less than 0.8 (but greater than 0.4, since otherwise there would be an additional loop depicted in the diagram). No other local extrema are apparent in the region of $\mathbf{R}^2$ shown.

Problem 5: Partial derivative practice

Compute the first and second partial derivatives in general, verifying equality of mixed partials directly, and evaluate the first partials at the indicated point $\mathbf{a}$.

(a) $g(x_1, x_2) = \sin(x_1x_2 - x_1 + x_2)$, $\mathbf{a} = (\sqrt{\pi}, \sqrt{\pi})$.

(b) $h(x, y) = e^x(x - y)^2$, $\mathbf{a} = (0, 1)$.

Solution:

(a) $g_{x_1} = (x_2 - 1)\cos(x_1x_2 - x_1 + x_2)$ and $g_{x_2} = (x_1 + 1)\cos(x_1x_2 - x_1 + x_2)$, so

$$g_{x_1}(\sqrt{\pi}, \sqrt{\pi}) = 1 - \sqrt{\pi}, \;\; g_{x_2}(\sqrt{\pi}, \sqrt{\pi}) = -1 - \sqrt{\pi}.$$

Also, $g_{x_1x_1} = -(x_2 - 1)^2 \sin(x_1x_2 - x_1 + x_2)$, $g_{x_2x_2} = -(x_1 + 1)^2 \sin(x_1x_2 - x_1 + x_2)$, and (in both ways) $g_{x_1x_2} = \cos(x_1x_2 - x_1 + x_2) - (x_2 - 1)(x_1 + 1)\sin(x_1x_2 - x_1 + x_2)$.

(b) $h_x = e^x(x - y)^2 + 2e^x(x - y)$ and $h_y = -2e^x(x - y)$, so

$$h_x(0, 1) = -1, \;\; h_y(0, 1) = 2.$$

Also, $h_{xx} = 2e^x(x-y) + e^x(x-y)^2 + 2e^x + 2e^x(x-y)$, $h_{yy} = 2e^x$, and (in both ways) $h_{xy} = -2e^x(x-y) - 2e^x$.

Problem 6: Finding candidates for local extrema

For each of the following functions $f : \mathbf{R}^2 \to \mathbf{R}$, find all critical points.

(a) $x_1^2 + 4x_1x_2 + 5x_2^2 - 4x_1 + 2x_2$.

(b) $x^4y^4 - 2x^2 - 2y^2$.

(c) $\cos(\pi(x^2 + y^2))$.

(d) $x_1^3 - 3x_1x_2^2 + 3x_2^2$.

Solution:

(a) We have $f_{x_1} = 2x_1 + 4x_2 - 4$ and $f_{x_2} = 4x_1 + 10x_2 + 2$. These simultaneously vanish only at the point $(12, -5)$.

(b) We have $f_x = 4x^3y^4 - 4x$ and $f_y = 4x^4y^3 - 4y$. These simultaneously vanish only at the five points $(0, 0)$, $(\pm 1, \pm 1)$.

(c) We have $f_x = -2x\pi \sin(\pi(x^2 + y^2))$ and $f_y = -2y\pi \sin(\pi(x^2 + y^2))$. These simultaneously vanish exactly when $x^2 + y^2$ is an integer since $\sin(\pi t) = 0$ precisely for t an integer, as we see by thinking about the meaning of such vanishing in terms of the unit circle and angles. (The case $x = y = 0$ is treated separately.)

(d) We have $f_{x_1} = 3x_1^2 - 3x_2^2$ and $f_{x_2} = -6x_1x_2 + 6x_2$. These simultaneously vanish only at the three points $(0,0), (1,\pm 1)$.

Problem 7: Computing extrema on a region I

Find the global extreme values of $f(x,y) = 2x^2 + y^2 + 5y$ on the disk of points (x,y) satisfying $x^2 + y^2 \leq 16$.

Solution: By computing the partial derivatives f_x and f_y to see where both vanish at the same point, the function f has one critical point on the interior of the disk: $(x,y) = (0,-5/2)$. On the boundary $x^2 + y^2 = 16$ the function can be expressed entirely in terms of y as

$$2(16 - y^2) + y^2 + 5y, \; -4 \leq y \leq 4$$

which by single-variable calculus has a critical point when its derivative $-4y + 2y + 5$ vanishes, which is to say $y = 5/2$ (this is in the open interval $(-4,4)$), in which case $x = \pm\sqrt{39}/2$ due to being on the boundary circle $x^2 + y^2 = 16$. Of course the interval's endpoints $y = \pm 4$ also must be considered, corresponding to $x = 0$.

So overall we have five candidates for where extrema are attained: $(0,-5/2)$, $(0,\pm 4)$, $(\pm\sqrt{39}/2, 5/2)$. Evaluating the function and seeing which values are largest and smallest, we get a global maximum $f(\pm\sqrt{39}/2, 5/2) = 153/4 = 38.25$ and global minimum $f(0,-5/2) = -25/4 = -6.25$. (The values of f at $(0,\pm 4)$ are $16 \pm 20 = -4, 36$.)

Problem 8: Computing extrema on a region II

Find the global extreme values of $f(x,y) = x^4y^4 - 2x^2 - 2y^2$ on the region of points (x,y) that lies on or inside the triangle with vertices $(-2,-2)$, $(-2,2)$, $(2,-2)$. (Sketch this triangle first, to get oriented.)

Solution: The function f has five critical points $(0,0)$, $(\pm 1, \pm 1)$ (seen by computing the partial derivatives f_x and f_y and determining they simultaneously vanish, which requires a bit of algebra). All but $(1,1)$ lie in the region (though only $(-1,-1)$ is on the interior). Next, we extremize on the boundary, by substitution on each boundary segment:

(i) On the segment from $(-2,-2)$ to $(-2,2)$ that is the part of the line $x = -2$ with $-2 \leq y \leq 2$, we substitute $x = -2$ to set up the task:

$$\text{extremize } 16y^4 - 8 - 2y^2 \text{ when } -2 \leq y \leq 2.$$

The derivative is $64y^3 - 4y$, which vanishes for $y = 0, \pm 1/4$. Accounting for the endpoints of this range of y-values, we get candidates $(-2,0)$, $(-2,\pm 1/4)$, $(-2,-2)$, $(-2,2)$ on this boundary segment at which an extreme value could be attained.

(ii) On the segment from $(-2,-2)$ to $(2,-2)$, which is the part of the line $y = -2$ with $-2 \leq x \leq 2$, we substitute $y = -2$ to set up the task:

$$\text{extremize } 16x^4 - 2x^2 - 8 \text{ when } -2 \leq x \leq 2$$

The derivative is $64x^3 - 4x$, which vanishes for $x = 0, \pm 1/4$. Accounting for endpoints of this range of x-values, we get the candidates $(0,-2)$, $(\pm 1/4, -2)$, $(-2,-2)$, $(2,-2)$ on this boundary segment at which an extreme value could be attained.

(iii) On the line segment from $(-2,2)$ to $(2,-2)$, which is the part of the line $y = -x$ with $-2 \leq x \leq 2$, we substitute $y = -x$ to set up the task:

$$\text{extremize } x^8 - 4x^2 \text{ when } -2 \leq x \leq 2$$

The derivative is $8x^7 - 8x$, which vanishes for $x = 0, \pm 1$. Accounting for the endpoints of this range of x-values, we get the candidates of $(0,0)$, $(1,-1)$, $(-1,1)$, $(-2,2)$, $(2,-2)$ on this boundary segment at which an extreme value could be attained.

Comparing values of the function of interest $f(x,y) = x^4y^4 - 2x^2 - 2y^2$ at all candidate points we have obtained, both the interior critical point $(-1,-1)$ and the various boundary points, we find a global minimum value $f\left(\pm\frac{1}{4}, -2\right) = f\left(-2, \pm\frac{1}{4}\right) = -8 - \frac{1}{16} \approx -8.06$ and a global maximum value $f(2,-2) = f(-2,2) = f(-2,-2) = 240$.