

Midterm 2 Review Solutions

1. Gaussian & Gauss-Jordan Elimination

Solve the following systems using Gaussian elimination.

1.

$$\begin{aligned}x + y + z &= 5 \\2x + 3y + 5z &= 8 \\4x + 5z &= 2\end{aligned}$$

Solution: We first need to put the equation into matrix-vector form:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \\ 2 \end{bmatrix}$$

We then form the augmented matrix so we can perform Gauss-Jordan elimination:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 2 & 3 & 5 & 8 \\ 4 & 0 & 5 & 2 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & -4 & 1 & -18 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 13 & -26 \end{array} \right]$$

Now that we have the matrix in upper-triangular form, we can easily solve via backwards substitution to find:

$$z = -2 \Rightarrow y = 4 \Rightarrow x = 3 \quad \blacksquare$$

2.

$$\begin{aligned}x + 2y - 3z &= 2 \\6x + 3y - 9z &= 6 \\7x + 14y - 21z &= 13\end{aligned}$$

Solution: Putting the system into matrix-vector form:

$$\begin{bmatrix} 1 & 2 & -3 \\ 6 & 3 & -9 \\ 7 & 14 & -21 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 13 \end{bmatrix}$$

Then, form the augmented matrix and perform Gauss-Jordan elimination:

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 6 & 3 & -9 & 6 \\ 7 & 14 & -21 & 13 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & -3 & 2 \\ 0 & -9 & 9 & -6 \\ 0 & 0 & 0 & -1 \end{array} \right]$$

The last row has all zeros equal to a non-zero, so right away we know there is no solution.

3.

$$\begin{aligned}A + B + 2C &= 1 \\2A - B + D &= -2 \\A - B - C - 2D &= 4 \\2A - B + 2C - D &= 0\end{aligned}$$

Solution: Putting the system into matrix-vector form:

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 \\ 1 & -1 & -1 & -2 \\ 2 & -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 4 \\ 0 \end{bmatrix}$$

Then forming the augmented matrix and performing row elimination:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 1 \\ 2 & -1 & 0 & 1 & -2 \\ 1 & -1 & -1 & -2 & 4 \\ 2 & -1 & 2 & -1 & 0 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 1 \\ 0 & -3 & -4 & 1 & -4 \\ 0 & -2 & -3 & -2 & 3 \\ 0 & -3 & -2 & -1 & -2 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 1 \\ 0 & -3 & -4 & 1 & -4 \\ 0 & 0 & -\frac{1}{3} & -\frac{8}{3} & \frac{17}{3} \\ 0 & 0 & 2 & -2 & 2 \end{array} \right] \Rightarrow$$

We can multiply through the third row by 3 to make the arithmetic a bit cleaner:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 1 \\ 0 & -3 & -4 & 1 & -4 \\ 0 & 0 & -1 & -8 & 17 \\ 0 & 0 & 2 & -2 & 2 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 1 \\ 0 & -3 & -4 & 1 & -4 \\ 0 & 0 & -1 & -8 & 17 \\ 0 & 0 & 0 & -18 & 36 \end{array} \right]$$

Then, we can solve via backwards substitution:

$$D = -2 \Rightarrow C = -1 \Rightarrow B = 2 \Rightarrow A = 1 \quad \blacksquare$$

2. Partial Derivatives

1. For exercises 1(a) and (b), verify that $w_{xy} = w_{yx}$.

(a) $w = \ln(2x + 3y)$

Solution:

$$\begin{aligned}\frac{\partial w}{\partial x} &= \frac{2}{2x+3y} \Rightarrow \frac{\partial^2 w}{\partial y \partial x} = -\frac{6}{(2x+3y)^2} \\ \frac{\partial w}{\partial y} &= \frac{3}{2x+3y} \Rightarrow \frac{\partial^2 w}{\partial x \partial y} = -\frac{6}{(2x+3y)^2} \quad \blacksquare\end{aligned}$$

(b) $w = x \sin y + y \sin x + xy$

Solution:

$$\begin{aligned}\frac{\partial w}{\partial x} &= \sin y + y \cos x + y \Rightarrow \frac{\partial^2 w}{\partial y \partial x} = \cos y + \cos x + 1 \\ \frac{\partial w}{\partial y} &= x \cos y + \sin x + x \Rightarrow \frac{\partial^2 w}{\partial x \partial y} = \cos y + \cos x + 1 \quad \blacksquare\end{aligned}$$

2. Now let's talk about something with meaning. If we stand on an ocean shore and take a snapshot of the waves, the picture shows a regular pattern of peaks and valleys in an instant of time. We see periodic vertical motion in space, with respect to distance. If we stand in the water, we can feel the rise and fall of the water as the waves go by. We see periodic vertical motion in time. In physics, this beautiful symmetry is expressed by the one-dimensional wave equation:

$$\frac{\partial^2 w}{\partial t^2} = c^2 \frac{\partial^2 w}{\partial x^2}$$

Show that the functions below satisfy the wave equation.

(a) $w = \cos(2x + 2ct)$

Solution: Taking the second partials in x and t :

$$\begin{aligned}\frac{\partial w}{\partial x} &= -2 \sin(2x + 2ct) \implies \frac{\partial^2 w}{\partial x^2} = -4 \cos(2x + 2ct) \\ \frac{\partial w}{\partial t} &= -2c \sin(2x + 2ct) \implies \frac{\partial^2 w}{\partial t^2} = -4c^2 \cos(2x + 2ct) \\ (-4c^2 \cos(2x + 2ct)) &= c^2 (-4 \cos(2x + 2ct)) \quad \blacksquare\end{aligned}$$

(b) $w = 5 \cos(3x + 3ct) + e^{x+ct}$

Solution: Taking the second partials in x and t :

$$\begin{aligned}\frac{\partial w}{\partial x} &= -15 \sin(3x + 3ct) + e^{x+ct} \implies \frac{\partial^2 w}{\partial x^2} = -45 \cos(3x + 3ct) + e^{x+ct} \\ \frac{\partial w}{\partial t} &= -15c \sin(3x + 3ct) + c e^{x+ct} \implies \frac{\partial^2 w}{\partial t^2} = -45c^2 \sin(3x + 3ct) + c^2 e^{x+ct} \\ (-45c^2 \sin(3x + 3ct) + c^2 e^{x+ct}) &= c^2 (-45 \cos(3x + 3ct) + e^{x+ct}) \quad \blacksquare\end{aligned}$$

Note: If you take CME 104, you will learn how to solve this equation for $w(x, t)$.

3. Chain Rule

1. Assuming that in the following equations y is a differentiable function of x , find dy/dx at the given point.

(a) $xe^y + \sin(xy) + y - \ln 2 = 0, \quad (0, \ln 2)$

Solution:

$$\begin{aligned}F(x, y) &= xe^y + \sin(xy) + y - \ln 2 = 0 \\ \frac{\partial}{\partial x} F(x, y) &= e^y + y \cos(xy) \\ \frac{\partial}{\partial y} F(x, y) &= xe^y + x \cos(xy) + 1 \\ \frac{dy}{dx} &= -\frac{F_x}{F_y} \\ &= -\frac{e^y + y \cos(xy)}{xe^y + x \cos(xy) + 1}\end{aligned}$$

$$\begin{aligned}\left.\frac{dy}{dx}\right|_{(0,\ln 2)} &= -\frac{2+\ln 2}{1} \\ &= -(2+\ln 2) \quad \blacksquare\end{aligned}$$

(b) $xy + y^2 - 3x - 3 = 0, \quad (-1, 1)$

Solution:

$$\begin{aligned}F(x, y) &= xy + y^2 - 3x - 3 = 0 \\ \frac{\partial}{\partial x}F(x, y) &= y - 3 \\ \frac{\partial}{\partial y}F(x, y) &= x + 2y \\ \frac{dy}{dx} &= -\frac{F_x}{F_y} \\ &= -\frac{y-3}{x+2y} \\ \left.\frac{dy}{dx}\right|_{(-1,1)} &= -\frac{-2}{1} \\ &= -2 \quad \blacksquare\end{aligned}$$

2. For the following function, find the values of dz/dx and dz/dy at the point specified.

$$xe^y + ye^z + 2\ln x - 2 - 3\ln 2 = 0, \quad (1, \ln 2, \ln 3)$$

Solution:

$$\begin{aligned}F(x, y, z) &= xe^y + ye^z + 2\ln x - 2 - 3\ln 2 = 0 \\ \frac{\partial}{\partial x}F(x, y) &= e^y + \frac{2}{x} \\ \frac{\partial}{\partial z}F(x, y) &= ye^z \\ \frac{dz}{dx} &= -\frac{F_x}{F_z} \\ &= -\frac{e^y + \frac{2}{x}}{ye^z} \\ \left.\frac{dz}{dx}\right|_{(1,\ln 2,\ln 3)} &= -\frac{2+2}{3\ln(2)} \\ &= -\frac{4}{3\ln(2)} \quad \blacksquare \\ \frac{\partial}{\partial y}F(x, y) &= xe^y + e^z \\ \frac{dz}{dy} &= -\frac{F_y}{F_z} \\ &= -\frac{xe^y + e^z}{ye^z}\end{aligned}$$

$$\begin{aligned}\left. \frac{dz}{dy} \right|_{(1, \ln 2, \ln 3)} &= -\frac{2+3}{3\ln(2)} \\ &= -\frac{5}{3\ln(2)} \quad \blacksquare\end{aligned}$$

4. Directional Derivatives

1. (a) Find the derivative of $f(x, y, z) = x^3 - xy^2 - z$ at $(1, 1, 0)$ in the direction of $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$.

Solution: First calculate the unit vector in the direction of \mathbf{v} :

$$|\mathbf{v}| = \sqrt{2^2 + (-3)^2 + 6^2} = 7 \implies \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{2}{7}\mathbf{i} - \frac{3}{7}\mathbf{j} + \frac{6}{7}\mathbf{k}$$

Then calculate the gradient of f at the given point:

$$\nabla f = \begin{bmatrix} 3x^2 - y^2 \\ -2xy \\ -1 \end{bmatrix} \implies \nabla f(1, 1, 0) = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$

Finally, we can find the directional derivative:

$$\nabla f \cdot \frac{\mathbf{v}}{|\mathbf{v}|} = 2\frac{2}{7} + 2\frac{3}{7} - \frac{6}{7} = \frac{4}{7} \quad \blacksquare$$

- (b) In what directions does f change most rapidly at $(1, 1, 0)$? What are the rates of change in these directions?

Solution: The gradient changes most rapidly in the direction of its gradient. Let \mathbf{u} be the unit vector in the direction of the gradient:

$$|\nabla f| = \sqrt{(2)^2 + (-2)^2 + (-1)^2} = 3 \implies \mathbf{u} = \frac{2}{3}\mathbf{i} - \frac{2}{3}\mathbf{j} + \frac{1}{3}\mathbf{k}$$

Then, we can calculate the derivative in this direction:

$$\nabla f \cdot \mathbf{u} = \frac{4}{3} + \frac{4}{3} + \frac{1}{3} = 3 \quad \blacksquare$$

2. Find the direction in which $h(x, y, z) = \ln(x^2 + y^2 - 1) + y + 6z$ increases and decreases most rapidly at $(1, 1, 0)$, then find the derivative in this direction.

Solution: First calculate the gradient:

$$\nabla h = \begin{bmatrix} \frac{2x}{x^2+y^2-1} \\ \frac{2y}{x^2+y^2-1} + 1 \\ 6 \end{bmatrix} \implies \nabla h(1, 1, 0) = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$$

The magnitude is:

$$|\nabla h| = \sqrt{2^2 + 3^2 + 6^2} = 7 \quad \blacksquare$$

So the derivative in this direction is 7.

3. In what directions is the derivative of $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$ at $(1, 1)$ equal to zero?

Solution: Recall that the gradient is always perpendicular to the level curves of a function. Therefore, if a vector is orthogonal to the gradient then we know the function is not changing in that direction. So, to solve this we need to figure out what the gradient is, then find a vector orthogonal to it:

$$\nabla f = \begin{bmatrix} \frac{4xy^2}{(x^2+y^2)^2} \\ \frac{-4x^2y}{(x^2+y^2)^2} \end{bmatrix}$$

Evaluating at the given point:

$$\nabla f(1, 1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

So, $u = i + j$ will be orthogonal to the gradient, so the derivative is zero in any direction parallel to u .

5. Gradients

Compute the gradient of the following functions at the given point.

1. $f(x, y, z) = 2z^3 - 2(x^2 + y^2)z + \tan^{-1}(xz)$, $(1, 1, 1)$

Solution: First compute the gradient vector:

$$\nabla f = \begin{bmatrix} -4xz + \frac{z}{1+x^2z^2} \\ -4yz \\ 6z^2 - 2(x^2 + y^2) + \frac{x}{1+x^2z^2} \end{bmatrix}$$

Then evaluate at the given point:

$$\nabla f(1, 1, 1) = \begin{bmatrix} -\frac{7}{2} \\ -4 \\ \frac{5}{2} \end{bmatrix} \quad \blacksquare$$

2. $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2} + \ln(xyz)$, $(-1, 2, -1)$

Solution:

$$\nabla f = \begin{bmatrix} -\frac{x}{(x^2+y^2+z^2)^{3/2}} + \frac{1}{x} \\ -\frac{y}{(x^2+y^2+z^2)^{3/2}} + \frac{1}{y} \\ -\frac{z}{(x^2+y^2+z^2)^{3/2}} + \frac{1}{z} \end{bmatrix}$$

Then evaluate at the given point:

$$\nabla f(-1, 2, -1) = \begin{bmatrix} \frac{1}{6^{3/2}} - 1 \\ -\frac{2}{6^{3/2}} + \frac{1}{2} \\ \frac{1}{6^{3/2}} - 1 \end{bmatrix} \quad \blacksquare$$

3. $f(x, y, z) = e^{x+y} \cos(z) + (y+1)\sin^{-1}(x)$, $(0, 0, \pi/6)$

Solution:

$$\nabla f = \begin{bmatrix} e^{x+y} \cos(z) + \frac{y+1}{\sqrt{1-x^2}} \\ e^{x+y} \cos(z) + \sin^{-1}(x) \\ -e^{x+y} \sin(z) \end{bmatrix}$$

Evaluating at the given point:

$$\nabla f(0, 0, \pi/6) = \begin{bmatrix} \frac{\sqrt{3}}{2} + 1 \\ \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix}$$

6. Double Integrals

1. Sketch the region of integration and reverse the order of integration. Then, evaluate the integral.

(a) $\int_0^{\pi/6} \int_{\sin(x)}^{1/2} x y^2 dy dx$

Solution: We can swap the order of integration as:

$$\int_0^{1/2} \int_0^{\sin^{-1}(y)} x y^2 dx dy$$

To evaluate, we know that the values are the same regardless of what order of integration we use, so in this case it is probably easiest to use the original form:

$$\begin{aligned} \int_0^{\pi/6} \int_{\sin(x)}^{1/2} x y^2 dy dx &= \int_0^{\pi/6} x \left. \frac{1}{3} y^3 \right|_{\sin(x)}^{1/2} dx \\ &= \int_0^{\pi/6} \left(\frac{x}{24} - \frac{x}{3} \sin^3(x) \right) dx = \frac{-200 + 36\sqrt{3}\pi + \pi^2}{1728} \quad \blacksquare \end{aligned}$$

(b) $\int_0^8 \int_{x^{1/3}}^2 \frac{1}{y^4 + 1} dy dx$

Solution: Rewriting the bounds of integration:

$$\int_0^2 \int_0^{y^3} \frac{1}{y^4 + 1} dx dy$$

Then solving:

$$\begin{aligned} \int_0^2 \int_0^{y^3} \frac{1}{y^4 + 1} dx dy &= \int_0^2 \frac{y^3}{y^4 + 1} dy \\ &= \left. \frac{1}{4} \ln(y^4 + 1) \right|_0^2 \\ &= \frac{1}{4} \ln(17) \quad \blacksquare \end{aligned}$$

2. Convert to polar coordinates, then evaluate the integral.

(a) $\int_0^4 \int_{-\sqrt{16-y^2}}^{\sqrt{16-y^2}} \sqrt{25-x^2-y^2} dx dy$

Solution: The region of integration is a semicircle of radius 4 in the top half-plane. So, we can write the integral in polar coordinates as:

$$\int_0^{\pi} \int_0^4 \sqrt{25-r^2} r dr d\theta$$

From here, it is straightforward to evaluate:

$$\begin{aligned}\int_0^\pi \int_0^4 \sqrt{25-r^2} r dr d\theta &= \int_0^\pi \left. -\frac{1}{3}(25-r^2)^{3/2} \right|_0^4 dr d\theta \\ &= \int_0^\pi \frac{98}{3} d\theta \\ &= \frac{98\pi}{3} \quad \blacksquare\end{aligned}$$

(b) $\int_{-1}^1 \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \ln(x^2 + y^2 + 1) dx dy$

Solution: The region of integration is a circle of radius 1 centered at the origin. From here, we can rewrite the integral in polar coordinates and solve:

$$\begin{aligned}\int_{-1}^1 \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \ln(x^2 + y^2 + 1) dx dy &= \int_0^{2\pi} \int_0^1 \ln(r^2 + 1) r dr d\theta \\ &= \int_0^{2\pi} \left. \frac{1}{2}((r^2 + 1)\ln(r^2 + 1) - r^2) \right|_0^1 d\theta \\ &= 2\pi \left(\ln(2) - \frac{1}{2} \right) \quad \blacksquare\end{aligned}$$

7. Unconstrained Optimization

1. Consider the function $f(x, y) = x^2 + kxy + y^2$. for what values of k will $f(x, y)$ have a saddle? A minimum? A maximum? Mathematically justify your answers.

Solution: Remember that it is the *second order conditions* that characterizes whether an extremum is a maximum, minimum, or saddle. For this problem, we first need to use the *first order conditions* to find the optimum points, then plug these into our second order conditions to find the values of k for which we have each type of extremum.

Using the first order conditions (i.e. taking the first partials and setting equal to zero):

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x + ky = 0 \\ x &= -\frac{ky}{2} \\ \frac{\partial f}{\partial y} &= 2y + kx = 0 \\ y^* &= 0 \implies x^* = 0\end{aligned}$$

From here, we need to solve for the second partials:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= 2 \\ \frac{\partial^2 f}{\partial y \partial x} &= k \\ \frac{\partial^2 f}{\partial y^2} &= 2\end{aligned}$$

The type of extremum is determined by the sign of the determinant of the Hessian $f_{xx}f_{yy} - f_{xy}^2$. (Remember that the Hessian is just the matrix of second partials.) Here, $f_{xx}f_{yy} - f_{xy}^2 = 4 - k^2$. Also, note that $f_{xx} > 0$, so we will either have a minimum or saddle point i.e. we cannot have a maximum no matter the value of k .

The condition for a minimum is $f_{xx}f_{yy} - f_{xy}^2 > 0$, so $-2 < k < 2$ gives a local minimum.

The condition for a saddle is $f_{xx}f_{yy} - f_{xy}^2 < 0$, so $|k| > 2$ gives a saddle.

The test is inconclusive for $k = \pm 2$.

2. Find the maxima and minima of $f(x, y) = 48xy - 32x^3 - 24y^2$ on $0 \leq x \leq 1$ and $0 \leq y \leq 1$.

Solution: Taking the first partials:

$$\frac{\partial f}{\partial x} = 48y - 96x^2 = 0$$

$$\frac{\partial f}{\partial y} = 48x - 48y = 0$$

$$x^* = y^*$$

$$x^* = 0, \frac{1}{2}$$

Taking the second partials (to determine the first order conditions):

$$\frac{\partial^2 f}{\partial x^2} = -192x$$

$$\frac{\partial^2 f}{\partial y^2} = -48$$

$$\frac{\partial^2 f}{\partial y \partial x} = 48$$

$f_{yy} < 0$ for all points and $f_{xx}f_{yy} - f_{xy}^2 > 0$, so we have a local maximum of 0 at (0,0) and local maximum of 2 at (1/2, 1/2).

We also need to check the boundaries for minima or maxima. For $x = 0$, we have $f(y) = -24y^2$ which has a minimum of -24 at $y = 1$. For $y = 0$, we have $f(x) = -32x^3$ which has a minimum of -32 at $x = 1$. For $x = 1$, $f(y) = 48y - 32 - 24y^2$ which has a minimum of -32 at $y = 1$. For $y = 1$, $f(x) = 48x - 32x^3 - 24$ which has a minimum of -24 at $x = 0$.

So, for the given domain we have a global maximum at (1/2, 1/2) and global minimum (1, 0).

3. Find local max, min, and saddle points of $f(x, y) = e^x(x^2 - y^2)$.

Solution: We first take the first partials to solve for the first order conditions:

$$\frac{\partial f}{\partial x} = 2xe^x + x^2e^x - y^2e^x = 0$$

$$\frac{\partial f}{\partial y} = -2ye^x = 0$$

$$y^* = 0$$

$$x^* = 0, -2$$

For the second partials:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= 2e^x + 2xe^x + 2xe^x + x^2e^x - y^2e^x \\ \frac{\partial^2 f}{\partial y \partial x} &= -2ye^x \\ \frac{\partial^2 f}{\partial y^2} &= -2e^x\end{aligned}$$

f_{yy} is strictly negative, so we will only have local maxima or saddles.

For $(0,0)$, we have $f_{xx}f_{yy} - f_{xy}^2 = 0(0) - 0 = 0$, so the test is inconclusive.

For $(-2,0)$, we have $f_{xx}f_{yy} - f_{xy}^2 = (2e^2 + 4e^2 + 4e^2)(0) - (-2e^2) > 0$, so we have a local maximum.

8. Constrained Optimization

- Find the points on the curve $x^2 + xy + y^2 = 1$ in the xy -plane that are nearest to and farthest from the origin.

Solution: Here the objective is to minimize or maximize Distance $= \sqrt{x^2 + y^2}$. Our method of Lagrange multipliers is agnostic to whether we are seeking a maximum or minimum, so we can proceed as usual and determine the two solutions:

We first setup the Lagrangian:

$$\mathcal{L} = \sqrt{x^2 + y^2} - \lambda(1 - x^2 - xy - y^2)$$

Taking the first partials with respect to x , y , and λ :

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= \frac{x}{\sqrt{x^2 + y^2}} + 2\lambda x + y = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= \frac{y}{\sqrt{x^2 + y^2}} + 2\lambda y + x = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= -1 + x^2 + xy + y^2 = 0\end{aligned}$$

We can rearrange the first two equation then divide them to find:

$$\frac{x}{y} = \frac{2\lambda x + y}{2\lambda y + x} \implies 2\lambda xy + x^2 = 2\lambda xy + y^2 \implies x^2 = y^2 \implies x = \pm y$$

For the first case $x = y$:

$$\begin{aligned}-1 + x^2 + x(x) + (x)^2 &= 0 \\ 3x^2 &= 1 \\ x^* &= \pm \frac{1}{\sqrt{3}} = y^*\end{aligned}$$

For the second case $x = -y$:

$$-1 + x^2 - x(x) + (-x)^2 = 0$$

$$x^2 = 1$$

$$x^* = \pm 1$$

$$y^* = \mp 1$$

So, we have minima at $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ and $(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$ and maxima at $(1, -1)$ and $(-1, 1)$.

2. Minimize the function $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraints $x + 2y + 3z = 6$ and $x + 3y + 9z = 9$.

Solution: Remember that when we have multiple constraints, we simply add an extra multiplier for each additional constraint. We setup the Lagrangian as:

$$\mathcal{L} = x^2 + y^2 + z^2 - \lambda_1(6 - x - 2y - 3z) - \lambda_2(9 - x - 3y - 9z)$$

Taking the first partials:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= 2x + \lambda_1 + \lambda_2 = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= 2y + 2\lambda_1 + 3\lambda_2 = 0 \\ \frac{\partial \mathcal{L}}{\partial z} &= 2z + 3\lambda_1 + 9\lambda_2 = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda_1} &= 6 - x - 2y - 3z = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda_2} &= 9 - x - 3y - 9z = 0\end{aligned}$$

Notice that we have five linear equations and five unknowns, so we can form the augmented matrix to easily solve for the five variables:

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 1 & 3 & 9 & 0 & 0 \\ 2 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 2 & 3 \\ 0 & 0 & 2 & 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

which after a lot of algebra, we find the solution is:

$$\begin{aligned}x &= \frac{81}{59} \\ y &= \frac{123}{59} \\ z &= \frac{9}{59}\end{aligned}$$

3. Let a_1, a_2, \dots, a_n be positive numbers. Find the maximum of $\sum_{i=1}^n a_i x_i$ subject to $\sum_{i=1}^n x_i^2 = 1$.

Solution: This problem is exactly the same as previous problems, except our objective and constraint are a bit more complicated to work with. Namely, we now have n variables to deal with

instead of the 2 or 3 that we're used to. However, you should get comfortable working with objectives/constraints of this type, since this is the form in which you will most often see optimization problems.

Set up the Lagrangian as usual:

$$\mathcal{L} = \sum_{i=1}^n a_i x_i - \lambda \left(1 - \sum_{i=1}^n x_i^2 \right)$$

It is useless to try to write out the first partial for every x_i , so instead we can write it out for a single x_i and our multiplier λ . Therefore, the first order conditions are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i} &= a_i + 2\lambda x_i = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 1 - \sum_{i=1}^n x_i^2 = 0 \end{aligned}$$

Rearrange the first equation and square to find:

$$a_i^2 = 4\lambda^2 x_i^2$$

Then, sum over all i for both sides to find:

$$\sum_{i=1}^n a_i^2 = 4\lambda^2 \sum_{i=1}^n x_i^2 = 4\lambda^2 \implies \lambda = \pm \frac{\sqrt{\sum_{i=1}^n a_i^2}}{2}$$

Using the first equation in our first order conditions:

$$a_i + 2 \left(\pm \frac{\sqrt{\sum_{i=1}^n a_i^2}}{2} \right) x_i = 0 \implies x_i = \pm \frac{a_i}{\sqrt{\sum_{i=1}^n a_i^2}}$$

We are looking for the maximum, so we should only consider the positive x_i . Therefore, the solution is:

$$x_i = \frac{a_i}{\sqrt{\sum_{i=1}^n a_i^2}} \quad \blacksquare$$

9. Linear Regression

1. A herpetologist has been tasked with studying the local salamander population. To do so, they have been collecting data on the average winter temperature (in °C) in a region and how many salamanders per square kilometer they find on a given patch of land the following spring. The data points are (8, 2), (6, 1), (11, 3), and (3, 1) (with salamander population give in hundreds of salamanders). They would like to model the relationship between the average temperature and salamander population, and believe a linear model may work well.

Solve for the least-squares best fit line for these four data points. What are the squared errors for each data point? Do you think a linear fit is appropriate?

Solution: We can use the least squares equations as given in the notes:

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

$$\begin{bmatrix} 230 & 28 \\ 28 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 58 \\ 7 \end{bmatrix}$$

Which has solution:

$$\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \frac{9}{34} \\ -\frac{7}{68} \end{bmatrix}$$

2. In some cases, the linear fit may not be very good. An alternative to fitting a line through data is to fit a higher-order polynomial such as a quadratic. That is, we fit a quadratic of the form:

$$y(x) = m_1 x^2 + m_2 x + b$$

to our data, which requires us to solve out for m_1 , m_2 , and b .

Reformulate the least-squares problem to fit a quadratic, and write out the system we should solve for m_1 , m_2 , and b .

Solution: The challenge here is re-deriving the squared error. Remember that the error is the difference between the estimated point and the actual data point, and the squared error is this term squared and summed over all points. So, we can write squared error for the quadratic as:

$$E = \sum_{i=1}^n (m_1 x_i^2 + m_2 x_i + b - y_i)^2$$

As an optimization problem, we are looking for the three coefficients that will minimize this. So, we can take the partial w.r.t. each of our three variables, set the partials to 0, and derive a system for m_1 , m_2 , and b that satisfy the first order conditions.

$$\frac{\partial E}{\partial m_1} = \sum_{i=1}^n 2x_i^2 (m_1 x_i^2 + m_2 x_i + b - y_i) = 0$$

$$\frac{\partial E}{\partial m_2} = \sum_{i=1}^n 2x_i (m_1 x_i^2 + m_2 x_i + b - y_i) = 0$$

$$\frac{\partial E}{\partial b} = \sum_{i=1}^n 2(m_1 x_i^2 + m_2 x_i + b - y_i) = 0$$

We can rewrite these equations as:

$$m_1 \sum_{i=1}^n x_i^4 + m_2 \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i$$

$$m_1 \sum_{i=1}^n x_i^3 + m_2 \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i$$

$$m_1 \sum_{i=1}^n x_i^2 + m_2 \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i$$

Which we can put into matrix form as:

$$\begin{bmatrix} \sum_{i=1}^n x_i^4 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i^2 y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

3. Sometimes an outlier can have a huge effect on our solution, or if we are fitting a high-degree polynomial, the coefficients may become very large and our model will *overfit* the data. (Indeed, this is one of the biggest problems that machine learning researchers face when formulating models.)

To combat this, we *regularize* the regression coefficients. Regularization involves adding a penalty term to our squared error to penalize large coefficients:

$$E = \sum_{i=1}^n (m x_i + b - y_i)^2 + \lambda(m^2 + b^2)$$

Rewrite the linear regression equations and include the regularization terms.

Solution: We follow the same steps as in the original derivation, with the added twist that we have the regularization term when we take the first derivatives.

$$\begin{aligned} \frac{\partial E}{\partial m} &= \sum_{i=1}^n 2x_i(m x_i + b - y_i) + 2\lambda m = 0 \\ \frac{\partial E}{\partial b} &= \sum_{i=1}^n 2(m x_i + b - y_i) + 2\lambda b = 0 \end{aligned}$$

We can rewrite these equations as:

$$\begin{aligned} m \left(\lambda + \sum_{i=1}^n x_i^2 \right) + b \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ m \sum_{i=1}^n x_i + b(\lambda + n) &= \sum_{i=1}^n y_i \end{aligned}$$

Which we can put into matrix form as:

$$\begin{bmatrix} \lambda + \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n + \lambda \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

10. MATLAB

1. For the function $f(x, y) = \tan^{-1}(x) + y^2$, write MATLAB code to compute $\int_1^{10} \int_2^4 f(x, y) dy dx$ using a right Riemann sum. Use 1001 grid points in each dimension.

Solution: First recall the formula for the right Riemann sum:

$$\int_{a_x}^{b_x} \int_{a_y}^{b_y} f(x, y) dy dx \approx \sum_{i=1}^{n_y} \sum_{j=1}^{n_x} f(x_i + \Delta x, y_j + \Delta y) \Delta x \Delta y$$

We can then implement this as:

```
[X Y] = meshgrid(linspace(1,10,1001), linspace(2,4,1001));
f = atan(X) + Y.^2;
dx = (10 - 1)/1000;
dy = (4 - 2)/1000;
I = sum(sum(f(2:end,2:end)))*dx*dy;
```

2. The steepest descent method is one of the cornerstones of optimization theory. In fact, steepest descent is the main workhorse for most machine learning algorithms in industry, and forms the backbone of most systems used in other fields. The steepest descent algorithm is given as follows:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k)$$

where x_k is your current point, γ_k is your step size at step k , and $\nabla f(x_k)$ is the gradient calculated at your current point.

- (a) Based on your knowledge of gradient vectors, what is the intuition behind this algorithm? Will it always converge to a local minimum? A global minimum? No need to mathematically justify your answers here—we are just looking for a qualitative explanation.

Solution: We know that the gradient always points in the steepest direction uphill. In this algorithm, we always move in the direction opposite the gradient. By doing so, we are guaranteed that the objective will always decrease.

- (b) Consider the function

$$f(x, y) = 2e^{-y}(x-2)^2 + e^{-x}(y-4)^4$$

Write the gradient descent update equations.

Solution: Here, we simply need to derive the gradient, then plug into the above update equation.

$$\begin{aligned}\frac{\partial f}{\partial x} &= 4e^{-y}(x-2) - e^{-x}(y-4)^4 \\ \frac{\partial f}{\partial y} &= -2e^{-y}(x-2)^2 + 4e^{-x}(y-4)^3\end{aligned}$$

- (c) For the function in the previous part, write MATLAB code that implements gradient descent with an initial guess of (0,0) and runs for 5,000 iterations. Use an initial step size of 10^{-2} and reduce the step size by 1/2 every 1000 iterations.

Solution:

```
df = @(x) [4*exp(-x(2))*(x(1) - 2) - exp(-x(1))*(x(2) - 4)^4; ...
          -2*exp(-x(2))*(x(1) - 2)^2 + 4*exp(-x(1))*(x(2) - 4)^3];
```

```
x = [0; 0];
gamma = 1e-2;
```

```
for i = 1:5000
    x = x - gamma*df(x);
    if mod(i,1000)==0
        gamma = 0.5*gamma;
    end
end
```