Chapter 5.

1. a. The equations $g(x, y, z) = 0$ and $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ define a system

$$\begin{aligned} x^2 + y^2 + z^2 &= 1 \\ 2 &= 2\lambda x \\ 3 &= 2\lambda y \\ -4 &= 2\lambda z. \end{aligned}$$

Multiplying the first equation by λ^2 and squaring the remaining equations produces

$$\begin{aligned} \lambda^2 x^2 + \lambda^2 y^2 + \lambda^2 z^2 &= \lambda^2 \\ 1 &= \lambda^2 x^2 \\ 9/4 &= \lambda^2 y^2 \\ 4 &= \lambda^2 z^2. \end{aligned}$$

Combining all these equations we find

$$1 + 9/4 + 4 = \lambda^2$$

and that tells us

$$\lambda = \pm\sqrt{29}/2.$$

At last, we solve for the remaining variables, finding

$$(x, y, z) = \pm\frac{1}{\sqrt{29}}(2, 3, -4).$$

That is, when $\lambda = \sqrt{29}/2$ we must take $(x, y, z) = (1/\sqrt{29})(2, 3, -4)$ and when $\lambda = -\sqrt{29}/2$ we must take $(x, y, z) = (1/\sqrt{29})(-2, -3, 4)$.

b. We want to find (x, y, z, λ) so that

$$\begin{aligned} \frac{\partial F}{\partial x} &= 2 + 2\lambda x = 0 \\ \frac{\partial F}{\partial y} &= 3 + 2\lambda y = 0 \\ \frac{\partial F}{\partial z} &= -4 + 2\lambda z = 0 \\ \frac{\partial F}{\partial \lambda} &= x^2 + y^2 + z^2 - 1 = 0. \end{aligned}$$

After rearrangement, these are essentially the same as the equations from part a., except that λ has been replaced by $-\lambda$. Using similar methods as before, the solution is found to be

$$(x, y, z, \lambda) = \pm\frac{1}{\sqrt{29}}\left(2, 3, -4, -\frac{29}{2}\right).$$

c. The solutions have identical values of x, y, and z, and opposite values of λ.

2. Let $g(x, y) = y$ so S is the curve $y = 0$, aka the x axis. Let $f(x, y) = y - x^3$ and take $M = 0$ so that the level curve L has equation $y = x^3$. This is evidently tangent to S at $(0, 0)$, which we take to be point p. Clearly L and S cross at p. This implies that p is not a local minimum or maximum of f restricted to the constraint curve S. To see this explicitly, observe that on S $f(x, y) = f(x, 0) = -x^3$ and that takes on values greater than $f(p) = 0$ to the left of p and less than $f(p)$ to the right of p. However, we can easily verify that p satisfies the Lagrange conditions. It certainly satisfies the equation $g(x, y) = 0$. And since $\nabla f(0, 0) = (0, 1) = \nabla g(0, 0)$, the gradients are parallel at $(0, 0)$. Thus we have shown in this example that a solution of the Lagrange conditions need not be either a local constrained maximum or minimum of f. More generally, this illustrates that Lagrange conditions are necessary, but not sufficient, for identifying local constrained maxima and minima.

3. Define $f(x, y, z) = x^2 + y^2 + z^2$ and $g(x, y, z) = xyz - 1$. On the constraint surface $g(x, y, z) = 0$ we observe that neither x nor y can vanish, and that $z = 1/xy$. Therefore, define $\Phi(x, y) = (x, y, 1/xy)$, so that the restriction of f to the constraint surface can be identified with $f \circ \Phi$.

   Now compute

$$d\Phi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{-1}{x^2y} & \frac{-1}{xy^2} \end{bmatrix}$$

   and notice that the columns are linearly independent. That is, they cannot be multiples of one another because no multiple of $(1, 0)$ equals $(0, 1)$.

   Now suppose (x^*, y^*, z^*) is the location of a constrained maximum or minimum of f. Then $z^* = 1/x^*y^*$, and $f \circ \Phi$ has a local max or min at (x^*, y^*). Thus, its gradient is equal to 0 there. But

$$\nabla(f \circ \Phi)(x^*, y^*) = (\nabla f(x^*, y^*, z^*))^T \cdot d\Phi(x^*, y^*)$$

so $\nabla f(x^*, y^*, z^*)$ is orthogonal to the columns of $d\Phi(x^*, y^*)$. Geometrically, $\nabla f(x^*, y^*, z^*)$ is perpendicular to the plane containing the columns of $d\Phi(x^*, y^*)$. By similar reasoning, since $g \circ \Phi$ is constant, its gradient also vanishes, and that implies $\nabla g(x^*, y^*, z^*)$ is also perpendicular to the same plane. Together, these observations imply that $\nabla f(x^*, y^*, z^*)$ and $\nabla g(x^*, y^*, z^*)$ are parallel.

4. The Lagrangian function in this example is $F(x, y, \lambda) = x^2 + y^2 + \lambda(xy - 1)$, and the solution to the constrained optimization occurs where $(x, y, \lambda) = (1, 1, -2)$. At that point we know that ∇F is zero. To apply the second derivative test we have to compute the Hessian matrix of second partial derivatives

$$H = \begin{bmatrix} 2 & \lambda & y \\ \lambda & 2 & x \\ y & x & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ -2 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

To show that $(1, 1, -2)$ is a saddle point, it suffices to show that there are both positive and negative eigenvalues. To that end, we find the characteristic polynomial

$$p(t) = \det \begin{bmatrix} t-2 & 2 & -1 \\ 2 & t-2 & -1 \\ -1 & -1 & t \end{bmatrix}.$$

Before computing the determinant, we observe that setting $t = 4$ makes the first two rows identical. Thus $p(4) = 0$ revealing one root of p (and a positive eigenvalue of H).

Proceeding, we expand the determinant formula for $p(t)$ and divide out a factor of $(t-4)$. This leads to $p(t) = (t-4)(t^2-2)$, and tells us that the eigenvalues of H are 4 and $\pm\sqrt{2}$. In particular, H has both positive and negative eigenvalues, as desired.

5. The n variable analog of the theorem is as follows.

Let f and g be functions of n variables, with continuous second derivatives. Let $F(x_1, \cdots, x_n, \lambda) = f(x_1, \cdots, x_n) + \lambda g(x_1, \cdots, x_n)$. Then, if $(x_1^*, \cdots, x_n^*, \lambda^*)$ is a critical point of F at which ∇g^* is not zero, F has a saddle point at $(x_1^*, \cdots, x_n^*, \lambda^*)$.

In the proof, we can again assume that $(x_1^*, \cdots, x_n^*) = 0$ and that ∇g^* points in the direction of the positive x_1 axis. Thus $\nabla g^* = (g_{x_1}^*, \cdots, g_{x_n}^*) = (a, 0, 0, \cdots, 0)$ for some $a \neq 0$.

Next we will consider the restriction of F to the two dimensional plane spanned by the vectors $(1, 0, 0, \cdots, 0)$ and $(0, 0, \cdots, 0, 1)$ in

$(x_1, \cdots, x_n, \lambda)$ space. This plane is characterized by the equations $x_k = 0$ for $2 \le k \le n-1$.

Let

$$h(x_1, \lambda) = F(x_1, 0, \cdots, 0, \lambda) = f(x_1, 0, \cdots, 0) + \lambda g(x_1, 0, \cdots, 0).$$

Then

$$\nabla h(x_1, \lambda) = (f_{x_1}(x_1, 0, \cdots, 0) + \lambda g_{x_1}(x_1, 0, \cdots, 0), g(x_1, 0, \cdots, 0)).$$

This vanishes at $(x_1, \lambda) = (0, \lambda^*)$, so $(0, \lambda^*)$ is a critical point of h.

We show that $(0, \lambda^*)$ is a saddle point of h by using the second derivative test for a function of two variables. To simplify the notation we will use x in place of x_1. We need the Hessian matrix of second partial derivatives, which is defined by

$$H(x, \lambda) = \begin{bmatrix} f_{xx}(x, 0, \cdots, 0) + \lambda g_{xx}(x, 0, \cdots, 0) & g_x(x, 0, \cdots, 0) \\ g_x(x, 0, \cdots, 0) & 0 \end{bmatrix}.$$

At the critical point, the determinant of the Hessian is $\det H(0, \lambda^*) = -(g_x^*)^2 = -a^2$, and this is negative because we know $a \ne 0$. Therefore, the second derivative test shows that h has a saddle point at $(0, \lambda^*)$. Hence F must have a saddle point at $(x_1^*, \cdots, x_n^*, \lambda^*)$, as claimed.

6. a. Substituting the parameteric expressions for x and y changes our problem to one of maximizing $\sqrt{2}(\cos t + \sin t)$ for $0 \le t \le 2\pi$. But that is the same as maximizing $2(\sqrt{1/2}\cos t + \sqrt{1/2}\sin t) = 2\sin(t + \pi/4)$. By inspection the maximum occurs when $t = \pi/4$ and that corresponds to the point $(x, y) = (1, 1)$.

   b. At the point $(1, 1)$ the directional derivative of F_λ in the direction of the vector $(-1, 1)$ is given by

   $$\sqrt{1/2}(-1, 1) \cdot \nabla F_\lambda(1, 1) = \sqrt{1/2}(-1, 1) \cdot (1+2\lambda, 1+2\lambda) = 0.$$

   c. At the point $(1, 1)$ the directional derivative of F_λ in the direction of the vector $(1, 1)$ is given by

   $$\sqrt{1/2}(1, 1) \cdot \nabla F_\lambda(1, 1) = \sqrt{1/2}(1, 1) \cdot (1+2\lambda, 1+2\lambda) = \sqrt{2}(1+2\lambda).$$

   As required, this vanishes if and only if $\lambda = -1/2$.

d. Since $F_{-1/2}(x, y) = x + y - (x^2 + y^2 - 2)/2$, we can compute directly $\nabla F_{-1/2}(x, y) = (1 - x, 1 - y)$. This shows that $(x, y) = (1, 1)$ is a critical point of $F_{-1/2}$

e. The results of the earlier parts of this problem show that in the family of functions F_λ, there is one for which $(1, 1)$ is a critical point. For that value of λ, the graph of F_λ has a horizontal tangent plane at the point $(1, 1)$. This is also where F_λ is maximized subject to the given constraint, because F_λ and f agree at each point of the constraint curve. Thus, this example illustrates the idea of leveling: it is possible to level the graph of f at the constrained optimum by adding the perturbation $(-1/2)g$, and therefore it is possible to find a λ so that the Lagrange conditions hold at the solution point.

Chapter 6.

1. The slope of the line from $(-a, a)$ to $(6-a, 6-a)$ is $(6-2a)/6 = (3-a)/3$. The point slope formula then gives the equation of the line as
$$y - a = \frac{3-a}{3}(x+a),$$
or equivalently
$$\frac{3-a}{3}(x+a) - y + a = 0.$$
Therefore, we define
$$F_a(x,y) = \frac{3-a}{3}(x+a) - y + a.$$
The equation $F_a(x,y) = 0$ then defines our family of lines. Applying the envelope algorithm, we differentiate with respect to a to find
$$\frac{\partial}{\partial a}F_a(x,y) = \frac{-1}{3}(x+a) + \frac{3-a}{3} + 1 = -\frac{1}{3}x - \frac{2}{3}a + 2.$$
Setting this to zero, we find that $a = -x/2 + 3$. Substituting in $F_a(x,y) = 0$ now gives the equation
$$\frac{x/2}{3}(x - x/2 + 3) - y - x/2 + 3 = 0$$
for the envelope of the family of lines. Algebraic rearrangement reduces this to $y = x^2/12 + 3$ which is immediately recognized as a parabola with vertex at $(0,3)$ and axis of symmetry along the y axis.

2. The intercepts of the kth line are $p_k = (N-k, 0)$ and $q_k = (0, k)$. This tells us that the kth line has equation
$$\frac{x}{N-k} + \frac{y}{k} = 1$$
or equivalently,
$$kx + (N-k)y = k(N-k).$$

This defines a family of lines, with k playing the role of parameter, and N held constant. To find the envelope we differentiate with respect to k. That gives

$$x - y = N - 2k$$

so on the envelope $k = (N-x+y)/2$. Then $N-k = (N+x-y)/2$. Substituting these in the equation for the family of lines now produces

$$2(N - x + y)x + 2(N + x - y)y = [N - (x - y)][N + (x - y)].$$

Now there are a couple of ways to proceed. Noticing the appearance of $x + y$ and $x - y$, we can introduce a new set of axes w and z by rotating the original axis through a 45° angle. The equation for the envelope can be expressed relative to the rotated axes using the substitutions $z = (x + y)/\sqrt{2}$ and $w = (x - y)/\sqrt{2}$, and that leads to an equation in w and z that is quadratic in one variable and linear in the other. This shows that the envelope is a parabola. Alternatively, we can simply expand the envelope equation in x and y and express it in the form $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$. This will necessarily be a conic section, and the condition for it to be a parabola is $B^2 - 4AC = 0$.

3. The solution to this problem is very similar to the preceding solution, and the details are left to the reader.

4. Take the origin at the intersection of the two lines. Express the endpoints of an initial string as nonzero vectors $\mathbf{u}$ and $\mathbf{v}$. Then the endpoints of the next string will be $(1 + \alpha)\mathbf{u}$ and $(1 - \beta)\mathbf{v}$ for some positive constants α and β. Since endpoints are equally spaced on each line, this shows that the end points of the kth line are $(1 + k\alpha)\mathbf{u}$ and $(1 - k\beta)\mathbf{v}$.

   At this point, it is possible to introduce coordinates for $\mathbf{u}$ and $\mathbf{v}$ and proceed as in the earlier problems. However, the algebra gets quite involved, and even with a computer algebra system, it can be quite a chore to verify that the envelope is a parabola. It is not difficult to see that the equation has to reduce to something of the form $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, but verifying that $B^2 - 4AC = 0$ is not necessarily trivial.

   An alternative is to show that a configuration of the sort already discussed can be linearly transformed into the current configuration. This transformation carries the lines of one string art design to the lines of the other, and because linear transformations preserve tangency, it also maps the envelope of the first design to the envelope of the second. To complete this line of argument, we need to know that a linear transformation always maps a parabola to a parabola, a familiar result from analytic geometry.

So, to carry out this program, consider the points $p_k = (1+\alpha k, 0)$ on the x axis and $q_k = (0, 1-\beta k)$ on the y axis. Using the sort of analysis applied in the two preceding problems, we can show that the envelope curve is of the form $Ax^2+Bxy+Cy^2+Dx+Ey+F = 0$ and that $B^2 - 4AC = 0$. In fact, I found that the envelope equation reduces to something of the form

$$(\beta x - \alpha y)^2 = Dx + Ey + F.$$

These results show that the envelope of the string art design defined by the points p_k and q_k is indeed a parabola.

Now consider the linear transformation that takes an arbitrary vector (x, y) to the vector $x\mathbf{u}+y\mathbf{v}$. In matrix notation, the transformation is given by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = M \begin{bmatrix} x \\ y \end{bmatrix}$$

where M is the 2×2 matrix with columns $\mathbf{u}$ and $\mathbf{v}$. Observe that $T(p_k) = (1 + \alpha k)\mathbf{u}$ and $T(q_k) = (1 - \beta k)\mathbf{v}$. This shows that T carries the string art design we just analyzed onto the original design for this problem. Therefore, in the original design, the envelope is always a parabola.

5. Let the roots be $x_1, \cdots, x_n$. Then $|a_0| = x_1 \cdots x_n$, and we want to maximize this product subject to the constraint that the sum of the roots is S. Using the proof in this chapter of the arithmetic mean - geometric mean inequality, the maximum value of the product occurs when all the roots are equal, say they all equal r. In this case, $S = nr$ so $r = S/n$, the polynomial is $(x - S/n)^n$, and the maximal value of $|a_0|$ is $(S/n)^n$.

6. Using the preceding problem, the polynomial is $(x - (n+1)/n)^n$, and the maximal value of $|a_0|$ is $(1 + 1/n)^n$. The limiting value is e.

7. Since x and y are non-negative, so is $xy = A$. This shows $0 \leq A$. For the other condition, observe that $xy \leq (x + y)^2/4$ if and only

if $4xy \leq (x+y)^2$, and that holds if and only if $0 \leq (x+y)^2 - 4xy$. But this last identically equals $x^2 - 2xy + y^2$, a perfect square. This shows that $xy \leq (x+y)^2/4$ and so $A \leq P^2/16$.