Here is an example of using a 1-constraint Lagrange multiplier technique to help solve a global max/min problem. Let

$$f(x, y) = 2xy - 2x^3y - 2xy^3$$

be our function. We located the critical points here and even did an example of find the absolute extrema over the unit square. Suppose however we are required to maximize or minimize f over the region

$$g(x, y) = x^4 + xy + y^4 \leqslant 3 \text{ in first quadrant.}$$

$$\nabla f(x, y) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle = \langle\, 2y - 6x^2y - 2y^3, 2x - 2x^3 - 6xy^2 \,\rangle$$

Since we cannot parameterize $x^4 + xy + y^4 = 3$, we turn to Lagrange multipliers.

$$\nabla g(x, y) = \left\langle \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right\rangle = \langle\, 4x^3 + y, x + 4y^3 \,\rangle$$

The Lagrange equations are

$$2y - 6x^2y - 2y^3 = \lambda(4x^3 + y) \tag{1}$$

$$2x - 2x^3 - 6xy^2 = \lambda(x + 4y^3) \tag{2}$$

$$(2y - 6x^2y - 2y^3)(x + 4y^3) = (2x - 2x^3 - 6xy^2)(4x^3 + y)$$

$$2xy - 6x^4y - 2xy^3 + 8y^3 - 24x^2y^4 - 8y^6 = 2yx - 6xy^4 - 2x^3y + 8x^3 - 24x^4y^2 - 8x^6$$

$$6(xy^4 - x^4y) + 2(x^3y - xy^3) - 8(x^3 - y^3) + 24(x^4y^2 - x^2y^4) + 8(x^6 - y^6) = 0$$

$$6xy(y^3 - x^3) + 2xy(x^3 - y^3) + 8(x^3 - y^3) + 24x^2y^2(x^2 - y^2) + 8(x^6 - y^6) = 0$$

$$(8 - 4xy)(x^3 - y^3) + 24x^2y^2(x^2 - y^2) + 8(x^6 - y^6) = 0$$

One solution is $x = y$ and you can check that in the region under consideration, these are the only solutions. When $x = y$ and $x^4 + xy + y^4 = 3$, $x = \pm 1$ and since we are in the first quadrant, $x = y = 1$. The value of $f(1, 1) = -2$.

We also need to check the lines from $(0, 0)$ to $(\sqrt[4]{3})$ and $(\sqrt[4]{3}, 0)$. The value of f is equal to 0 along both these lines. Hence the maximum value of f occurs at the interior local maximum, $\left(\frac{1}{2}, \frac{1}{2}\right)$. The minimum value occurs at $(1, 1)$ where the value is -2.

Here is an example of a 2-constraint dimension 3 Lagrange multiplier problem.
Find the extrema (maxima and minima) of $x^2 + y^2 + z^2$ subject to the constraints

$$x + y - z = 1 \tag{3}$$

$$z^2 = x^2 + y^2 \tag{4}$$

Let $f(x, y, z) = x^2 + y^2 + z^2$, $g(x, y, z) = x + y - z$ and $h(x, y, z) = z^2 - x^2 - y^2$.
Then

$$\nabla f = \langle 2x, 2y, 2z\rangle \tag{5}$$

$$\nabla g = \langle 1, 1, -1\rangle \tag{6}$$

$$\nabla h = \langle -2x, -2y, 2z\rangle \tag{7}$$

Then Lagrange says to solve

$$\langle 2x, 2y, 2z\rangle = \lambda\langle 1, 1, -1\rangle + \mu\langle -2x, -2y, 2z\rangle$$

or

$$2x = \lambda - 2x\mu \tag{8}$$

$$2y = \lambda - 2y\mu \tag{9}$$

$$2z = -\lambda + 2z\mu \tag{10}$$

or

$$\lambda = 2x(1 + \mu) \tag{11}$$

$$\lambda = 2y(1 + \mu) \tag{12}$$

$$\lambda = 2z(-1 + \mu) \tag{13}$$

If $\mu \neq -1$, $x = y$ and so $z = \pm\sqrt{2}\,x$. Moreover, $2x - (\pm\sqrt{2}\,x) = 1$ so $x = \dfrac{1}{2-\sqrt{2}}$, $y = \dfrac{1}{2-\sqrt{2}}$ and $z = \dfrac{\sqrt{2}}{2-\sqrt{2}}$. No need to simplify but if you must, $x = \dfrac{2+\sqrt{2}}{2}$, $y = \dfrac{2+\sqrt{2}}{2}$ and $z = \dfrac{2+2\sqrt{2}}{2} = 1+\sqrt{2}$.

The other point can be worked out so the two critical points are

$$\left(\frac{2+\sqrt{2}}{2}, \frac{2+\sqrt{2}}{2}, 1+\sqrt{2}\right) \qquad \text{and} \qquad \left(\frac{2-\sqrt{2}}{2}, \frac{2-\sqrt{2}}{2}, 1-\sqrt{2}\right)$$

If $\mu = -1$, (11), (12) and (13) say $\lambda = 0$ and hence $z = 0$. But then $x^2 + y^2 = 0$ and so $x = y = 0$. But the point $(0, 0, 0)$ does not satisfy the first constraint equation, so there are only two critical points.

$$f\left(\frac{2+\sqrt{2}}{2}, \frac{2+\sqrt{2}}{2}, 1+\sqrt{2}\right) = 6 + 4\sqrt{2} \quad \text{and} \quad f\left(\frac{2-\sqrt{2}}{2}, \frac{2-\sqrt{2}}{2}, 1-\sqrt{2}\right) = 6 - 4\sqrt{2}.$$

Are these really maxima and/or minima?

We did not consider this question either in class or in the book but it is an important question. We know from our theory that the two points above are critical points for the curve which is the intersection of $x+y-z=1$ and $z^2=x^2+y^2$ but we know from theory that a critical point can be a local maximum, a local minimum, or neither and even if they are local extrema, they may not be global extrema.

In the case above, the intersection is an ellipse, a parabola or a hyperbola, depending on the plane. If it's an ellipse then there will be at least one global maxima and one global minima, but if it is a hyperbola there can be two local minima and no maxima or one of each. Which case are we in and what can we do if we can't narrow down our options until we can list the possibilities?

At least we can take a shot at the local max/min question via the second derivative test. Consider the function $f\big(\vec{r}(t)\big)$ and suppose $\vec{r}(0)$ is a critical point. Then the second derivative can be computed. Let us write the variables as x_1, x_2 and x_3 and write $\vec{r}(t) = \big(r_1(t), r_2(t), r_3(t)\big)$.

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2} = \frac{d\Big((\nabla f)\big(\vec{r}(t)\big) \bullet \vec{r}\,'(t)\Big)}{dt} = (\nabla f)\big(\vec{r}(t)\big) \bullet \vec{r}\,''(t) + \sum_{i=1}^{3} \frac{d\frac{\partial f}{\partial x_i}(\vec{r}(t))}{dt} \cdot r_i'(t)$$
$$= (\nabla f)\big(\vec{r}(t)\big) \bullet \vec{r}\,''(t) + \sum_{i=1}^{3} \Big(\nabla \frac{\partial f}{\partial x_i}(\vec{r}(t)) \bullet \vec{r}\,'(t)\Big) \cdot r_i'(t)$$

Theoretically we can parameterize the curve so that $\vec{r}\,'(t) = \nabla g\big(\vec{r}(t)\big) \times \nabla h\big(\vec{r}(t)\big)$.

We will need the Hessians of all three of our functions so for any function F, denote

$$\mathcal{H}(F) = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}$$

Then $\dfrac{d\frac{\partial F}{\partial x_i}\big(\vec{r}(t)\big)}{dt} = \nabla \dfrac{\partial F}{\partial x_i}\big(\vec{r}(t)\big) \bullet \vec{r}\,'(t) = \displaystyle\sum_{j=1}^{k} \frac{\partial^2 F}{\partial x_j \partial x_i}\big(\vec{r}(t)\big) \cdot r_j(t)$ where $\vec{r}\,'(t) = \nabla g\big(\vec{r}(t)\big) \times \nabla h\big(\vec{r}(t)\big) = \big(r_1(t), r_2(t), r_3(t)\big)$.

Given the Hessian $\mathcal{H}(F)$ and a vector $\vec{v} = (v_1, v_2, v_3)$ define

$$\mathcal{H}(F) * \vec{v} = \left\langle \sum_{i=1}^{3} F_{1i} v_i, \sum_{i=1}^{3} F_{2i} v_i, \sum_{i=1}^{3} F_{3i} v_i \right\rangle$$

Then

$$\frac{dF}{dt}\big(\vec{r}(t)\big) = \Big(\mathcal{H}(F)\big(\vec{r}(t)\big)\Big) * \vec{r}\,'(t)$$

Now

$$\vec{r}\,''(t) = \frac{d\nabla g\big(\vec{r}(t)\big) \times \nabla h\big(\vec{r}(t)\big)}{dt} = \frac{d\nabla g\big(\vec{r}(t)\big)}{dt} \times \nabla h\big(\vec{r}(t)\big) + \nabla g\big(\vec{r}(t)\big) \times \frac{d\nabla h\big(\vec{r}(t)\big)}{dt}$$
$$= \big(\mathcal{H}(g) * (\nabla g \times \nabla h)\big) \times \nabla h + \nabla g \times \big(\mathcal{H}(h) * (\nabla g \times \nabla h)\big)$$

Therefore, if $\vec{r}(0) = \vec{a} = (a_1, a_2, a_3)$,

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2}(\vec{a}\,) = \nabla f(\vec{a}\,) \bullet \Big(\big(\mathcal{H}(g) * (\nabla g(\vec{a}\,) \times \nabla h(\vec{a}\,))\big) \times \nabla h(\vec{a}\,)\Big) +$$
$$\nabla f(\vec{a}\,) \bullet \Big(\nabla g(\vec{a}\,) \times \big(\mathcal{H}(h)(\vec{a}\,) * (\nabla g(\vec{a}\,) \times \nabla h(\vec{a}\,))\big)\Big) +$$
$$\Big(\mathcal{H}(f)(\vec{a}\,) * \big((\nabla g(\vec{a}\,) \times \nabla h(\vec{a}\,)\big)\Big) \bullet (\nabla g(\vec{a}\,) \times \nabla h(\vec{a}\,))$$

$$\mathcal{L}(f)(\vec{a}) = \mathcal{H}(f)(\vec{a}) * \big((\nabla g(\vec{a}) \times \nabla h(\vec{a})\big) \tag{14}$$

$$\mathcal{L}(g)(\vec{a}) = \mathcal{H}(g)(\vec{a}) * \big((\nabla g(\vec{a}) \times \nabla h(\vec{a})\big) \tag{15}$$

$$\mathcal{L}(h)(\vec{a}) = \mathcal{H}(h)(\vec{a}) * \big((\nabla g(\vec{a}) \times \nabla h(\vec{a})\big) \tag{16}$$

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2}(\vec{a}) = \nabla f(\vec{a})\bullet\Big(\mathcal{L}(g)(\vec{a})\times\nabla h(\vec{a})\Big)+\nabla f(\vec{a})\bullet\Big(\nabla g(\vec{a})\times\mathcal{L}(h)(\vec{a})\Big)+\mathcal{L}(f)(\vec{a})\bullet\Big(\nabla g(\vec{a})\times\nabla h(\vec{a})\Big)$$

$$(A) \qquad \frac{d^2 f\big(\vec{r}(t)\big)}{dt^2}(\vec{a}) = \det\begin{pmatrix}\nabla f(\vec{a})\\ \mathcal{L}(g)(\vec{a})\\ \nabla h(\vec{a})\end{pmatrix} + \det\begin{pmatrix}\nabla f(\vec{a})\\ \nabla g(\vec{a})\\ \mathcal{L}(h)(\vec{a})\end{pmatrix} + \det\begin{pmatrix}\mathcal{L}(f)(\vec{a})\\ \nabla g(\vec{a})\\ \nabla h(\vec{a})\end{pmatrix}$$

No one can remember a formula like (A) for long. When you take a linear algebra course you will learn to write formulae like this more efficiently so they are easier to use. You may eventually take a differential geometry course where you will develop a real feel for what is going on and so (A) seems natural to you.

Anyway, in our case

$\nabla f = \langle 2x, 2y, 2z\rangle$.

$$\mathcal{H}(f) = \begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}$$

$\nabla g = \langle 1, 1, -1\rangle$,

$$\mathcal{H}(g) = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$\nabla h = \langle -2x, -2y, 2z\rangle$

$$\mathcal{H}(h) = \begin{pmatrix}-2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 2\end{pmatrix}$$

We need the cross product $\nabla g \times \nabla h$:

$$\det\begin{pmatrix}i & j & k\\ 1 & 1 & -1\\ -2x & -2y & 2z\end{pmatrix} = \big\langle 2(z-y), 2(x-z), 2(x-y)\big\rangle$$

$$\mathcal{L}(f) = \big\langle 4(z-y), 4(x-z), 4(x-y)\big\rangle \tag{17}$$

$$\mathcal{L}(g) = \big\langle 0, 0, 0\big\rangle \tag{18}$$

$$\mathcal{L}(h) = \big\langle -4(z-y), -4(x-z), 4(x-y)\big\rangle \tag{19}$$

$$\det\begin{pmatrix} 2x & 2y & 2z \\ 0 & 0 & 0 \\ -2x & -2y & 2z \end{pmatrix} + \det\begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & -1 \\ 4(y-z) & 4(z-x) & 4(x-y) \end{pmatrix} + \det\begin{pmatrix} 4(z-y) & 4(x-z) & 4(x-y) \\ 1 & 1 & -1 \\ -2x & -2y & 2z \end{pmatrix}$$

$$= 0 + \det\begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & -1 \\ 4(y-z) & 4(z-x) & 4(x-y) \end{pmatrix} - \det\begin{pmatrix} -2x & -2y & 2z \\ 1 & 1 & -1 \\ 4(z-y) & 4(x-z) & 4(x-y) \end{pmatrix}$$

$$= \Big(8x(z-y) + 8y(z-x) + 8z(2z-x-y)\Big) - \Big(-8x(2x-y-z) + 8y(x+z-2y) + 8z(x+y-2z)\Big)$$

$$= 16x(x-y) + 16y(y-x) + 16z(2z-x-y) = 16(x-y)^2 + 16z(2z-x-y)$$

so finally

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2} = 16\Big((x-y)^2 + z(2z-x-y)\Big)$$

Since $x+y-z=1$ and since at a critical point $x=y$, we have: *at a critical point*,

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2} = 16z(z-1)$$

Look at the critical point $\left(\frac{2+\sqrt{2}}{2}, \frac{2+\sqrt{2}}{2}, 1+\sqrt{2}\right)$.

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2} = 16 \cdot \big(1+\sqrt{2}\big) \cdot \sqrt{2} = 16\big(2+\sqrt{2}\big)$$

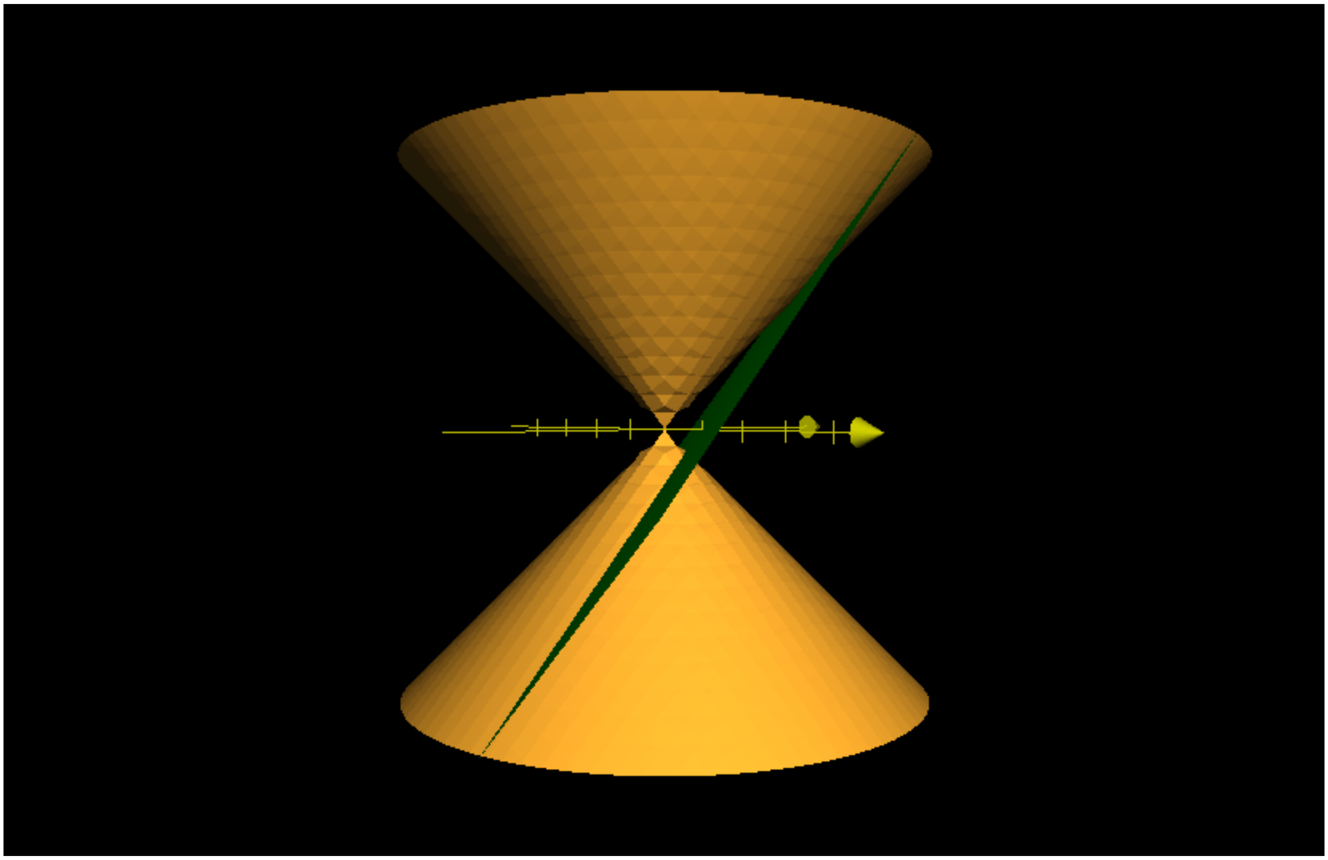
Hence the second derivative at this critical point is positive and so by the one variable second derivative test, this point is a local minimum!

At the critical point $\left(\frac{2-\sqrt{2}}{2}, \frac{2-\sqrt{2}}{2}, 1-\sqrt{2}\right)$.

$$\frac{d^2 f\big(\vec{r}(t)\big)}{dt^2} = 16\big(2-\sqrt{2}\big) > 0$$

Hence this point is also a local minimum!

Therefore, the intersection curve must be a hyperbola! I was wrong in class when I said it was an ellipse.



Here is a picture of the situation. The green object is the plane $x + y - z = 1$ and the yellow object is the cone $z^2 = x^2 + y^2$. The function being minimized is the distance^2 to the origin. The intersection is a hyperbola as can be seen by the different view below.