Optimization and Lagrange Multipliers

This material corresponds roughly to sections 14.7 and 14.8 in the book.

Taylor's Formula:
Let $T(x,y,z)$ be a scalar field and consider a small displacement $\triangle \mathbf{r} = (\triangle x, \triangle y, \triangle z)$. Suppose T is of class C^3, that is, T has partial derivatives of at least third order, each of them being continuous. Then **Taylor's formula** is

$$\begin{aligned} T\left(x+\triangle x, y+\triangle y, z+\triangle z\right) = T(x,y,z) &+ \frac{\partial T(x,y,z)}{\partial x}\triangle x + \frac{\partial T(x,y,z)}{\partial y}\triangle y + \frac{\partial T(x,y,z)}{\partial z}\triangle z \\ &+ \frac{1}{2}\left(\frac{\partial^2 T(x,y,z)}{\partial x^2}(\triangle x)^2 + \frac{\partial^2 T(x,y,z)}{\partial y^2}(\triangle y)^2 + \frac{\partial^2 T(x,y,z)}{\partial z^2}(\triangle z)^2\right) \\ &+ \frac{1}{2}\left(2\frac{\partial^2 T(x,y)}{\partial x \partial y}\triangle x \triangle y + 2\frac{\partial^2 T(x,y)}{\partial x \partial z}\triangle x \triangle z + 2\frac{\partial^2 T(x,y)}{\partial y \partial z}\triangle y \triangle z\right) + R_2(\triangle \mathbf{r}) \end{aligned}$$

where

$$\lim_{\triangle \mathbf{r} \longrightarrow \mathbf{0}} \frac{R_2\left(\triangle \mathbf{r}\right)}{\left|\triangle \mathbf{r}\right|^2} = 0 \tag{1}$$

More succinctly, we can write

$$T(\mathbf{r} + \triangle \mathbf{r}) = T(\mathbf{r}) + \nabla T\left(\mathbf{r}\right)\triangle \mathbf{r} + \frac{1}{2}\left(\triangle \mathbf{r}\right)^T H_T\left(\mathbf{r}\right)\triangle \mathbf{r} + R_2\left(\triangle \mathbf{r}\right) \tag{2}$$

Where we have represented $\triangle \mathbf{r}$ as a column vector (instead of a row vector) and $H_T\left(\mathbf{r}\right)$ is the Hessian **Hessian** of the scalar field

$$H_T = \begin{pmatrix} \frac{\partial^2 T}{\partial x^2} & \frac{\partial^2 T}{\partial x \partial y} & \frac{\partial^2 T}{\partial x \partial z} \\ \frac{\partial^2 T}{\partial x \partial y} & \frac{\partial^2 T}{\partial y^2} & \frac{\partial^2 T}{\partial y \partial z} \\ \frac{\partial^2 T}{\partial x \partial z} & \frac{\partial^2 T}{\partial y \partial z} & \frac{\partial^2 T}{\partial z^2} \end{pmatrix} \tag{3}$$

For those acquainted with linear algebra, notice that H_T is a symmetric matrix.

Critical points and Relative Extrema of a function $T(x,y)$:

- A **critical point** of T is a point (a,b) in the domain of T such that

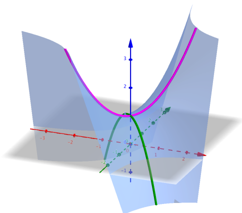
$$\frac{\partial T}{\partial x}(a,b) = 0 \quad \text{and} \quad \frac{\partial T}{\partial y}(a,b) = 0 \tag{4}$$

  that is, $\nabla T(a,b) = \mathbf{0}$, or at least one of the partial derivatives does not exist.

- T has a **relative maximum** at (a,b) if $T(x,y) \le T(a,b)$ for all points (x,y) that are sufficiently close to (a,b). We say $T(a,b)$ is a **relative maximum value**.

- T has a **relative minimum** at (a,b) with **relative minimum value** $T(a,b)$ if $T(x,y) \ge T(a,b)$ for all points (x,y) that are sufficiently close to (a,b)

- (a,b) is called a **saddle point** if it is a critical point but it is neither a relative minimum nor a relative maximum. For example, the origin $(0,0)$ is a saddle point for $f(x,y) = xy + 1$



The Second Derivative Test: to classify the relative extrema of a function $T(x,y)$

- Find the critical points of $T(x,y)$ by solving the system of equations

$$\frac{\partial T}{\partial x} = 0 \quad \text{and} \quad \frac{\partial T}{\partial y} = 0 \tag{5}$$

- Define the **discriminant**

$$D(x,y) = T_{xx}T_{yy} - T_{xy}^2 = \det H_T = \det \begin{pmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{pmatrix} \tag{6}$$

Let (a,b) be a critical point of T.

1. If $D(a,b) > 0$ and $T_{xx}(a,b) < 0$, $T(x,y)$ has a relative maximum at (a,b). [☹]
2. If $D(a,b) > 0$ and $T_{xx}(a,b) > 0$, $T(x,y)$ has a relative minimum at (a,b). [☺]
3. If $D(a,b) < 0$ then (a,b) is a saddle point.
4. If $D(a,b) = 0$ the test is inconclusive.

Problem 1. Classify the critical points of $f(x,y) = x^2y + \frac{1}{3}y^3 - x^2 - y^2 + 2$ using the second derivative test.

We find the critical points of f. The partial derivatives of f are

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy - 2x = 2x(y-1) \\ \frac{\partial f}{\partial y} &= x^2 + y^2 - 2y = x^2 + y(y-2) \end{aligned} \tag{7}$$

The critical points must solve the equations $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$, that is,

$$\begin{cases} 2x(y-1) = 0 \\ x^2 + y(y-2) = 0 \end{cases} \tag{8}$$

The first equation has solution $x = 0$ or $y = 1$. If we substitute $x = 0$ into the second equation we have

$$y(y-2) = 0 \implies y = 0 \text{ or } y = 2 \tag{9}$$

and so the two critical points corresponding to $x = 0$ are

$$(0,0), \quad (0,2) \tag{10}$$

If $y = 1$ we substitute it into the second equation to obtain

$$x^2 - 1 = 0 \implies x = \pm 1 \tag{11}$$

and so the corresponding critical points are

$$(1,1), \quad (-1,1) \tag{12}$$

Now we proceed to classify these critical points. To compute the discriminant we compute the second order derivatives

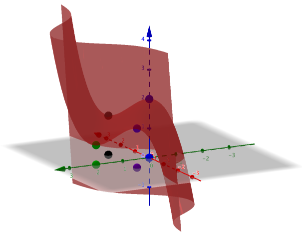
$$\begin{aligned} f_{xx} &= \frac{\partial}{\partial x}\left(2x(y-1)\right) = 2(y-1) \\ f_{xy} &= \frac{\partial}{\partial y}\left(2x(y-1)\right) = 2x \\ f_{yy} &= \frac{\partial}{\partial y}\left(x^2 + y^2 - 2y\right) = 2y - 2 = 2(y-1) \end{aligned} \tag{13}$$

The discriminant therefore is

$$\begin{aligned} D(x,y) &= f_{xx}f_{yy} - f_{xy}^2 \\ &= 4(y-1)^2 - 4x^2 \\ &= 4\left((y-1)^2 - x^2\right) \end{aligned} \tag{14}$$

We evaluate the discriminant at each critical point and apply the second derivative test:

Critical Point	$f_{xx}(x,y) = 2(y-1)$	$D(x,y) = 4\left((y-1)^2 - x^2\right)$	Classification
$(0,0)$	-2	4	relative maximum
$(0,2)$	2	4	relative minimum
$(1,1)$	0	-4	saddle point
$(-1,1)$	0	-4	saddle point



Problem 2. Show that the surface $z = xy$ has neither a maximum nor a minimum point.

The partial derivatives are

$$\frac{\partial z}{\partial x} = y \qquad \frac{\partial z}{\partial y} = x \tag{15}$$

and so the only critical point is the origin, that is, $(0,0)$. To show that it is a saddle point we compute $D(0,0)$. The second order partial derivatives are

$$\frac{\partial^2 z}{\partial x^2} = 0 \qquad \frac{\partial^2 z}{\partial x \partial y} = 1 \qquad \frac{\partial^2 z}{\partial y^2} = 0 \tag{16}$$

and so

$$D(x,y) = f_{xx} f_{yy} - f_{xy}^2 = -1 \tag{17}$$

In particular $D(0,0) = -1$ which implies by the second derivative test that $(0,0)$ is a saddle point.

Problem 3. If the product of the sines of the angles of a triangle is a maximum, show that the triangle is equilateral.

Call α, β, γ the three angles of the triangle. The product of the sines of the angles is

$$\sin\alpha \sin\beta \sin\gamma \tag{18}$$

and since these are the angles of a triangle

$$\alpha + \beta + \gamma = \pi \tag{19}$$

which means that we can solve for one the angles in terms of the other two

$$\gamma = \pi - \alpha - \beta \tag{20}$$

and so the function that we are trying to maximize is

$$\boxed{f(\alpha,\beta) = \sin\alpha \sin\beta \sin(\pi - \alpha - \beta)} \tag{21}$$

Since α, β are the angles of a triangle, we can assume that $0 < \alpha < \pi$ and $0 < \beta < \pi$. The partial derivatives of f are

$$\begin{aligned}
\frac{\partial f}{\partial \alpha} &= \cos\alpha \sin\beta \sin(\pi - \alpha - \beta) - \sin\alpha \sin\beta \cos(\pi - \alpha - \beta) \\
&= \sin\beta \left(\cos\alpha \sin(\pi - (\alpha+\beta)) - \sin\alpha \cos(\pi - (\alpha+\beta))\right) \\
\frac{\partial f}{\partial \beta} &= \sin\alpha \cos\beta \sin(\pi - \alpha - \beta) - \sin\alpha \sin\beta \cos(\pi - \alpha - \beta) \\
&= \sin\alpha \left(\cos\beta \sin(\pi - (\alpha+\beta)) - \sin\beta \cos(\pi - (\alpha+\beta))\right)
\end{aligned} \tag{22}$$

Before finding the critical points, we also compute the higher order derivatives

$$\begin{aligned}
\frac{\partial^2 f}{\partial \alpha^2} &= -2\sin\beta \left(\sin\alpha \sin(\pi - (\alpha+\beta)) + \cos\alpha \cos(\pi - (\alpha+\beta))\right) \\
\frac{\partial}{\partial \alpha}\left(\frac{\partial f}{\partial \beta}\right) &= \cos\alpha \left(\cos\beta \sin(\pi - (\alpha+\beta)) - \sin\beta \cos(\pi - (\alpha+\beta))\right) \\
&\quad - \sin\alpha \left(\cos\beta \cos(\pi - (\alpha+\beta)) + \sin\beta \sin(\pi - (\alpha+\beta))\right) \\
\frac{\partial^2 f}{\partial \beta^2} &= -2\sin\alpha (\sin\beta \sin(\pi - (\alpha+\beta)) + \cos\beta \cos(\pi - (\alpha+\beta)))
\end{aligned} \tag{23}$$

The critical points must solve the equations $\frac{\partial f}{\partial \alpha} = \frac{\partial f}{\partial \beta} = 0$, that is,

$$\begin{aligned}
\sin\beta \left(\cos\alpha \sin(\pi - (\alpha+\beta)) - \sin\alpha \cos(\pi - (\alpha+\beta))\right) = 0 \\
\sin\alpha \left(\cos\beta \sin(\pi - (\alpha+\beta)) - \sin\beta \cos(\pi - (\alpha+\beta))\right) = 0
\end{aligned} \tag{24}$$

Because we are assuming that $0 < \alpha < \pi$ and $0 < \beta < \pi$, $\sin\alpha$ and $\sin\beta$ are never 0, so we must solve the equations

$$\begin{aligned}
\cos\alpha \sin(\pi - (\alpha+\beta)) - \sin\alpha \cos(\pi - (\alpha+\beta)) = 0 \\
\cos\beta \sin(\pi - (\alpha+\beta)) - \sin\beta \cos(\pi - (\alpha+\beta)) = 0
\end{aligned} \tag{25}$$

If we use the identities $\sin(\theta_1 - \theta_2) = \sin\theta_1 \cos\theta_2 - \cos\theta_1 \sin\theta_2$ and $\cos(\theta_1 - \theta_2) = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2$ and so the equations are the same as

$$\boxed{\begin{aligned}
\cos\alpha \sin(\alpha+\beta) + \sin\alpha \cos(\alpha+\beta) = 0 \quad (\bullet) \\
\cos\beta \sin(\alpha+\beta) + \sin\beta \cos(\alpha+\beta) = 0 \quad (\bullet\bullet)
\end{aligned}} \tag{26}$$

Multiply the first equation by $\cos\beta$ and the second equation by $\cos\alpha$ to get

$$\begin{aligned} \cos\beta\cos\alpha\sin(\alpha+\beta)+\cos\beta\sin\alpha\cos(\alpha+\beta)=0 \quad (\star) \\ \cos\alpha\cos\beta\sin(\alpha+\beta)+\cos\alpha\sin\beta\cos(\alpha+\beta)=0 \quad (\star\star) \end{aligned} \tag{27}$$

If we subtract both equations, that is , $(\star)-(\star\star)$ to obtain

$$\begin{aligned} & \cos\beta\sin\alpha\cos(\alpha+\beta)-\cos\alpha\sin\beta\cos(\alpha+\beta)=0 \\ \implies & (\cos\beta\sin\alpha-\cos\alpha\sin\beta)\cos(\alpha+\beta)=0 \\ \implies & \sin(\alpha-\beta)\cos(\alpha+\beta)=0 \end{aligned} \tag{28}$$

Therefore, either $\sin(\alpha-\beta)=0$ or $\cos(\alpha+\beta)=0$. If $\cos(\alpha+\beta)=0$ then $\alpha+\beta=\frac{\pi}{2}$ and equations $(\bullet)$ and $(\bullet\bullet)$ become

$$\begin{aligned} \cos\alpha\sin(\alpha+\beta)=0 \implies \cos\alpha=0 \implies \alpha=\tfrac{\pi}{2} \\ \cos\beta\sin(\alpha+\beta)=0 \implies \cos\beta=0 \implies \beta=\tfrac{\pi}{2} \end{aligned} \tag{29}$$

Clearly each these equations can't be satisfied simultaneously so the case $\cos(\alpha+\beta)=0$ does not occur. The case $\sin(\alpha-\beta)=0$ implies that $\alpha=\beta$ and so $(\bullet)$ and $(\bullet\bullet)$ become the same equal to

$$\cos\alpha\sin(2\alpha)+\sin\alpha\cos(2\alpha)=0 \tag{30}$$

Because $\sin(2\alpha)=2\sin\alpha\cos\alpha$ and $\cos(2\alpha)=\cos^2\alpha-\sin^2\alpha$ we have the equation

$$\begin{aligned} & 2\sin\alpha\cos^2\alpha+\sin\alpha\cos^2\alpha-\sin^3\alpha=0 \\ \implies & 3\cos^2\alpha-\sin^2\alpha=0 \\ \implies & 3-3\sin^2\alpha-\sin^2\alpha=0 \\ \implies & 3=4\sin^2\alpha \\ \implies & \tfrac{3}{4}=\sin^2\alpha \\ \implies & \sin\alpha=\pm\tfrac{\sqrt{3}}{2} \end{aligned} \tag{31}$$

and since $0<\alpha<\beta$ we must have $\sin\alpha=\frac{\sqrt{3}}{2}$ which implies that $\alpha=\frac{\pi}{3}$. Therefore we found that

$$\alpha=\beta=\gamma=\frac{\pi}{3} \tag{32}$$

and so the triangle must be equilateral. To show that it maximizes $f(\alpha,\beta)$ we compute

$$f_{xx}\left(\frac{\pi}{3},\frac{\pi}{3}\right)=-2\left(\frac{\sqrt{3}}{2}\right)\left(\left(\frac{\sqrt{3}}{2}\right)^2+\left(\frac{1}{2}\right)^2\right) \tag{33}$$

which is negative. Similarly

$$\begin{aligned} D\left(\tfrac{\pi}{3},\tfrac{\pi}{3}\right) &= f_{xx}\left(\tfrac{\pi}{3},\tfrac{\pi}{3}\right) f_{yy}\left(\tfrac{\pi}{3},\tfrac{\pi}{3}\right) - \left(f_{xy}\left(\tfrac{\pi}{3},\tfrac{\pi}{3}\right)\right)^2 \\ &= 4\left(\tfrac{\sqrt{3}}{2}\right)^2\left(\left(\tfrac{\sqrt{3}}{2}\right)^2+\left(\tfrac{1}{2}\right)^2\right)^2 - \left(-\left(\tfrac{\sqrt{3}}{2}\right)\left(\tfrac{1}{2}\right)^2-\left(\tfrac{\sqrt{3}}{2}\right)^3\right) \end{aligned} \tag{34}$$

which is positive. Therefore, the second derivative test shows that we get a maximum, which is what we wanted to show.

Problem 4. Find the dimensions of a box (top included) which contains a given volume V and uses minimum material (i.e, has minimum surface area)

Let x, y, z be the sides of the box. The volume is

$$V = xyz \tag{35}$$

and the surface area is

$$S = 2xy + 2xz + 2yz \tag{36}$$

Since V is fixed we can solve for z and find

$$\boxed{z = \frac{V}{xy}} \tag{37}$$

Substituting in the formula for S we find the function we want to minimize

$$\boxed{S(x,y) = 2xy + 2\frac{V}{y} + 2\frac{V}{x}} \tag{38}$$

The partial derivatives are

$$\begin{aligned} \tfrac{\partial S}{\partial x} &= 2y - 2\tfrac{V}{x^2} \\ \tfrac{\partial S}{\partial y} &= 2x - 2\tfrac{V}{y^2} \end{aligned} \tag{39}$$

and the second order derivatives are

$$\begin{aligned} \tfrac{\partial^2 S}{\partial x^2} &= 4\tfrac{V}{x^3} \\ \tfrac{\partial}{\partial y}\left(\tfrac{\partial S}{\partial x}\right) &= 2 \\ \tfrac{\partial^2 S}{\partial y^2} &= 4\tfrac{V}{y^3} \end{aligned} \tag{40}$$

The critical points must satisfy the equations

$$\begin{aligned} 2y - 2\tfrac{V}{x^2} = 0 &\implies y = \tfrac{V}{x^2} \implies V = yx^2(\bullet) \\ 2x - 2\tfrac{V}{y^2} = 0 &\implies x = \tfrac{V}{y^2} \implies V = xy^2(\bullet\bullet) \end{aligned} \tag{41}$$

Setting $(\bullet) = (\bullet\bullet)$ we get $x = y$ and substituting in $(\bullet)$ we $V = x^3$ and so we have the box must be a cube of sides $x = \sqrt[3]{V}$. To show that it corresponds to a minimum observe that

$$S_{xx}\left(\sqrt[3]{V}, \sqrt[3]{V}\right) = 4 \tag{42}$$

and that

$$\begin{aligned} D(x,y) &= 16\tfrac{V^2}{x^3y^3} - 2 \\ \implies \quad D\left(\sqrt[3]{V}, \sqrt[3]{V}\right) &= 14 \end{aligned} \tag{43}$$

and so by the second derivative test we obtain a relative minimum.

Example 5. Classify the critical points of $f(x,y) = e^{2x+3y}\left(8x^2 - 6xy + 3y^2\right)$.

First we compute the partial derivatives of f

$$\frac{\partial f}{\partial x} = 2e^{2x+3y}\left(8x^2 - 6xy + 3y^2\right) + e^{2x+3y}\left(16x - 6y\right) \tag{44}$$

$$\frac{\partial f}{\partial y} = 3e^{2x+3y}\left(8x^2 - 6xy + 3y^2\right) + e^{2x+3y}\left(-6x + 6y\right) \tag{45}$$

The critical points must have vanishing partial derivatives, in other words, they must solve the system

$$\begin{cases} 2e^{2x+3y}\left(8x^2 - 6xy + 3y^2\right) + e^{2x+3y}\left(16x - 6y\right) = 0 \\ 3e^{2x+3y}\left(8x^2 - 6xy + 3y^2\right) + e^{2x+3y}\left(-6x + 6y\right) = 0 \end{cases} \tag{46}$$

Which is equivalent to the equations

$$\begin{cases} 8x^2 - 6xy + 3y^2 + 8x - 3y = 0 \\ 8x^2 - 6xy + 3y^2 - 2x + 2y = 0 \end{cases} \tag{47}$$

Subtracting both equations we find

$$10x = 5y \tag{48}$$

That is

$$2x = y \tag{49}$$

Now we substitute back in the first equation to find

$$8x^2 - 12x^2 + 12x^2 + 8x - 6x = 0 \tag{50}$$

Notice that this can be factorized as

$$x\left(8x + 2\right) = 0 \tag{51}$$

which means that either $x = 0$ or $x = -\frac{1}{4}$. Since $2x = y$ the critical points are $(0,0)$ and $\left(-\frac{1}{4}, -\frac{1}{2}\right)$.

To find the Hessian from

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 8x - 3y\right) \\ \frac{\partial f}{\partial y} &= 3e^{2x+3y}\left(8x^2 - 6xy + 3y^2 - 2x + 2y\right) \end{aligned} \tag{52}$$

we can compute the second order derivatives

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= 4e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 8x - 3y\right) + 4e^{2x+3y}\left(8x - 3y + 4\right) \\ \frac{\partial^2 f}{\partial x \partial y} &= 6e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 8x - 3y\right) + 6e^{2x+3y}\left(-2x + 2y - 1\right) \\ \frac{\partial^2 f}{\partial y^2} &= 9e^{2x+3y}\left(8x^2 - 6xy + 3y^2 - 2x + 2y\right) + 9e^{2x+3y}\left(-2x + 2y + \tfrac{2}{3}\right) \end{aligned} \tag{53}$$

So the Hessian is

$$H_f(x, y) = \begin{pmatrix} 4e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 16x - 6y + 4\right) & 6e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 6x - y - 1\right) \\ 6e^{2x+3y}\left(8x^2 - 6xy + 3y^2 + 6x - y - 1\right) & 9e^{2x+3y}\left(8x^2 - 6xy + 3y^2 - 4x + 4y + \tfrac{2}{3}\right) \end{pmatrix} \tag{54}$$

Evaluating at the critical point $(0, 0)$ we have

$$H_f\left(0, 0\right) = \begin{pmatrix} 16 & -6 \\ -6 & 6 \end{pmatrix} \tag{55}$$

Since $\det H_f(0, 0) = 60$ and the first entry is positive we conclude that $(0, 0)$ is a relative minimum.

For the critical point $\left(-\frac{1}{4}, -\frac{1}{2}\right)$

$$H_f\left(-\frac{1}{4}, -\frac{1}{2}\right) = \begin{pmatrix} 4e^{-2}\left(\frac{7}{2}\right) & 6e^{-2}\left(-\frac{3}{2}\right) \\ 6e^{-2}\left(-\frac{3}{2}\right) & 9e^{-2}\left(\frac{1}{6}\right) \end{pmatrix} \tag{56}$$

Since $\det H_f(-\frac{1}{4}, -\frac{1}{2}) = e^{-4}\left(-60\right) < 0$ we find that this corresponds to a saddle point.

Example 6. Find and classify the absolute maxima and minima of the function $f(x, y) = x^2 - xy + y^2 + 3x - 2y + 1$ defined on the rectangle $-2 \leq x \leq 0$**,** $0 \leq y \leq 1$

Whenever a continuous scalar field is defined on a bounded region which is also closed, it will achieve an absolute maximum and absolute minimum, similar to what happened for functions of one variable defined on a closed interval.

To find these in our case, we work first on the interior of the rectangle, where we can find the critical points in the usual way, and then work with the four boundary pieces of the rectangle separately.

As usual

$$\frac{\partial f}{\partial x} = 2x - y + 3 \qquad \frac{\partial f}{\partial y} = -x + 2y - 2 \tag{57}$$

so the critical points must solve

$$\begin{cases} 2x - y + 3 = 0 \\ -x + 2y - 2 = 0 \end{cases} \tag{58}$$

The previous system has solution $\left(-\frac{4}{3}, \frac{1}{3}\right)$, which does belong to the rectangle. Given

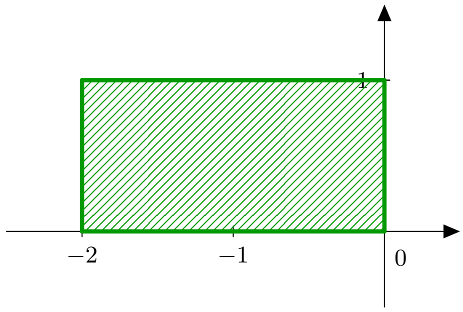


Figure 1: Optimization region

that

$$\frac{\partial^2 f}{\partial x^2} = 2 \qquad \frac{\partial^2 f}{\partial x \partial y} = -1 \qquad \frac{\partial^2 f}{\partial y^2} = 2 \tag{59}$$

the Hessian of the function is

$$H_f\left(-\frac{4}{3}, \frac{1}{3}\right) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \tag{60}$$

Since $\det H_f = 3$ and the first entry is negative this critical point is a relative minimun. It will also be useful to notice that $f\left(-\frac{4}{3}, \frac{1}{3}\right) = -\frac{4}{3}$

Now we analyze each side of the rectangle separately.

1. Side $-2 \leq x \leq 0 \quad y = 0$: Here $f(x, 0) = x^2 + 3x + 1$ and now we have turned the problem into a single variable function defined on a closed interval $[-2, 0]$. Just as we did when we were working with the rectangle, we analyze f at the endpoints $-2, 0$ and on the open interval $(-2, 0)$. For the endpoints notice that $f(-2, 0) = -1$ and $f(0, 0) = 1$. On the interval $(-2, 0)$ we have $\frac{\partial f(x,0)}{\partial x} = 2x + 3$ so the point $x = -\frac{3}{2}, y = 0$ is a candidate for an absolute maximum or minimum for the function. Notice that $f\left(-\frac{3}{2}, 0\right) = -\frac{5}{4}$

2. Side $x = 0 \quad 0 \leq y \leq 1$: Here $f(0, y) = y^2 - 2y + 1$. On the endpoints we have $f(0, 0) = 1$ and $f(0, 1) = 0$. For the interval $(0, 1)$, since $\frac{\partial f(0,y)}{\partial y} = 2y - 2$ we find the critical point $(0, 1)$, which we just computed.

3. Side $-2 \leq x \leq 0 \quad y = 1$: Here $f(x, 1) = x^2 + 2x$. Again $f(-2, 1) = 0$ and $\frac{\partial f(x,1)}{\partial x} = 2x + 2$ so there is a critical point $(-1, 1)$ and $f(-1, 1) = -1$

4. Side $x = -2 \quad 0 \leq y \leq 1$: Here $f(-2, y) = y^2 - 1$. Since $\frac{\partial f(-2,y)}{\partial y} = 2y$ we find the critical point $(-2, 0)$ which had already analyzed.

Therefore, our list of candidates for maxima and minima are the points (with corresponding values)

(x,y)	$f(x,y)$
$\left(-\frac{4}{3},\frac{1}{3}\right)$	$-\frac{4}{3}$
$(-2,0)$	-1
$(0,0)$	1
$\left(-\frac{3}{2},0\right)$	$-\frac{5}{4}$
$(0,1)$	0
$(-2,1)$	0
$(-1,1)$	-1

From here we can see that $(0,0)$ corresponds to the absolute maximum while the absolute minimum coincides with the relative minimum and happens at the point $\left(-\frac{4}{3},\frac{1}{3}\right)$.

Optimization with and without constraints.

So far we have been dealing with functions of two variables $f(x,y)$. However, our previous discussion carries easily to the case of functions of more variables. For example, if we have a function $f(x,y,z)$, the critical points would be found by solving the system of equations

$$\frac{\partial f}{\partial x}=0 \quad , \quad \frac{\partial f}{\partial y}=0, \quad \frac{\partial f}{\partial z}=0 \tag{61}$$

In other words, we are interested in finding the points $P=(a,b,c)$ where

$$\nabla f(P)=\mathbf{0}=(0,0,0) \tag{62}$$

Notice that for such a point P, all the directional derivatives vanish, that is,

$$D_{\mathbf{v}}f(P)=\underbrace{\nabla f(P)}_{\mathbf{0}}\cdot\mathbf{v}=\mathbf{0}\cdot\mathbf{v}=0 \tag{63}$$

Therefore, we can make the following definition of a critical point for a scalar field f:

Critical point of a scalar field (unconstrained case)
Let $f(x,y,z)$ denote a scalar field [for example, $f(x,y,z)$ could represent the temperature at the point (x,y,z)]. Then we say that $P=(a,b,c)$ is a critical point of f if the directional derivatives $D_{\mathbf{v}}f$ of f at P vanish in *all* possible directions, that is,

$$D_{\mathbf{v}}f(P)=0 \text{ for all direction vectors } \mathbf{v} \tag{64}$$

This occurs if and only if $\nabla f(P)=\mathbf{0}$, so we recover the original definition of critical point as being one where all partial derivatives vanish.

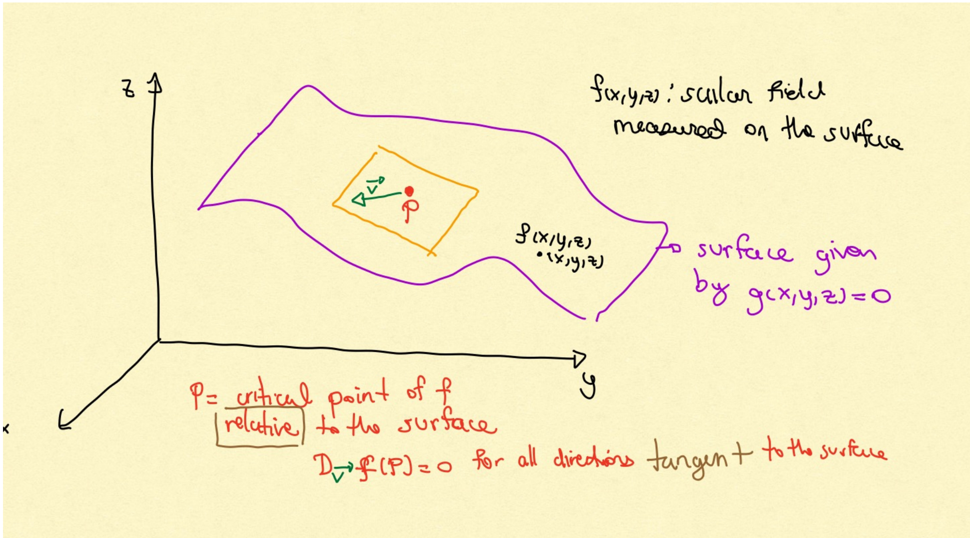
Now suppose you are measuring a scalar field f, but you are no longer allowed to move anywhere you want in space. More concretely, you could be trying to measure the values of f on the surface of a planet. Remember that we think of surfaces as being given by an equation of the form $g(x,y,z)=0$. For example,

$$g(x,y,z)=x^2+y^2+z^2-R^2=0 \tag{65}$$

represents the sphere of radius R centered at the origin, which is a surface.

In this new setup, we are interested in determining the local maxima and minima of f, but relative to the points on the surface determined by the equation $g(x, y, z) = 0$. In other words, if we say that the north pole is a relative minimum for the temperature f, what we mean is that compared to all nearby points *on the surface of the Earth*, then the north pole is a relative minimum. We are no longer interested in comparing the temperature at the north pole against the temperature at points in outer space.

This has implications for what the definition of a critical point of f should be *relative* to the surface $g(x, y, z) = 0$. The reason there is a difference is that we are no longer allowed to move freely in space. In other words, you can only move in ways which are *tangent* to the surface of the Earth.



Remember that at each point of the surface there is a tangent plane, which encodes precisely all the direction vectors which are tangent to the surface at that point. Therefore, we modify our definition of critical point as follows:

Critical point of a scalar field relative to some surface (constrained case)

Let $f(x, y, z)$ denote a scalar field [for example, $f(x, y, z)$ could represent the temperature at the point (x, y, z)]. Suppose that $g(x, y, z) = 0$ represents the equation of a surface and you measure f only at the points which belong to this surface.

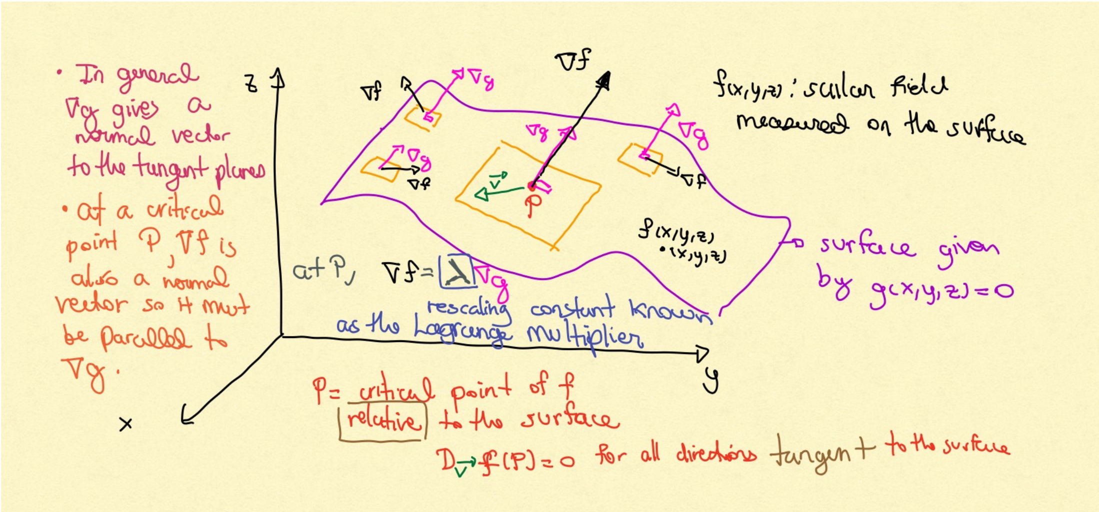
Then we say that $P = (a, b, c)$ **is a critical point of** f **relative to the surface** g if P is a point on the surface and all the directional derivatives $D_{\mathbf{v}}f$ of f at P vanish in *all tangent directions to the surface*, that is,

$$D_{\mathbf{v}}f(P) = 0 \text{ for all direction vectors } \mathbf{v} \text{ which belong to the tangent plane at } P \quad (66)$$

Notice that if $\mathbf{v}$ is a (unit) vector which belongs to the tangent plane to the surface through point P, then as before we have that

$$D_{\mathbf{v}}(P) = 0 \text{ is the same as } \nabla f(P) \cdot \mathbf{v} = 0 \quad (67)$$

So $\nabla f(P)$ must be perpendicular (orthogonal) to all the tangent vectors to the surface. But this is precisely the property the normal vector n to a plane satisfies. Moreover, we already know a vector with such a property. Namely, the gradient ∇g of the equation defining the surface is a normal vector to the tangent planes. This observation is the main idea behind the Lagrange Multipliers method.



Key idea behind Lagrange Multiplier's Method:
Let $f(x, y, z)$ denote a scalar field [for example, $f(x, y, z)$ could represent the temperature at the point (x, y, z)]. Suppose that $g(x, y, z) = 0$ represents the equation of a surface and you measure f only at the points which belong to this surface.
If $P = (a, b, c)$ is a critical point of f relative to the surface g, then at the point P the vector $\nabla f(P)$ is a normal vector for the tangent plane to the surface passing through P. Therefore, it must be parallel to $\nabla g(P)$, which always is a normal vector. In other words

$$\nabla f(P) = \lambda \nabla g \quad \text{for a critical point } P \tag{68}$$

The constant λ is a rescaling factor and it is known as the **Lagrange multiplier.** We think of f as the function we want to optimize, and g as the constraint equation.

Optimization with constraints: Lagrange's multiplier method
To find the critical points of the function $f(x, y, z)$ subject to the constraint $g(x, y, z) = 0$ we must solve the system of equations

$$\begin{cases} \nabla f = \lambda \nabla g \\ g(x, y, z) = 0 \end{cases} \tag{69}$$

When the function $f(x, y, z)$ is subject to the constraints $g(x, y, z) = 0$ and $h(x, y, z) = 0$ we must solve instead

$$\begin{cases} \nabla f = \lambda \nabla g + \mu \nabla h \\ g(x, y, z) = 0 \\ h(x, y, z) = 0 \end{cases} \tag{70}$$

Example 7. Find the critical points of $f(x, y, z) = 2x + 3y + z$ subject to the constraint $g(x, y, z) = 4x^2 + 3y^2 + z^2 - 80$

First we compute the gradients of f and g

$$\nabla f = 2\mathbf{i} + 3\mathbf{j} + \mathbf{k} \tag{71}$$

$$\nabla g = 8x\mathbf{i} + 6y\mathbf{j} + 2z\mathbf{k} \tag{72}$$

From equation 69 we must solve the system

$$\begin{cases} 2 = \lambda 8x \\ 3 = \lambda 6y \\ 1 = \lambda 2z \\ 4x^2 + 3y^2 + z^2 = 80 \end{cases} \tag{73}$$

In order to have a solution we clearly need $\lambda \neq 0$ in which case we find

$$x = \frac{1}{4\lambda} \quad y = \frac{1}{2\lambda} \quad z = \frac{1}{2\lambda} \tag{74}$$

substituting in the last equation

$$\frac{1}{4\lambda^2} + \frac{3}{4\lambda^2} + \frac{1}{4\lambda^2} = 80 \tag{75}$$

which is the same as

$$4\lambda^2 = \frac{5}{80} \tag{76}$$

so the values for λ are

$$\lambda = \pm\frac{1}{8} \tag{77}$$

and the critical points are $(2, 4, 4)$ and $(-2, -4, -4)$.

Example 8. Find the maximum volume of a box of rectangular base that must be inside the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

In this case we must maximize $V(x, y, z) = 8xyz$ subject to the constaint $g(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$. Using the Lagrange multipliers we must solve $\nabla V = \lambda \nabla g$, that is,

$$(8yz, 8xz, 8xy) = \lambda \left(\frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2} \right) \tag{78}$$

From the system of equations

$$\begin{cases} 4yz = \frac{\lambda x}{a^2} \\ 4xz = \frac{\lambda y}{b^2} \\ 4xy = \frac{\lambda z}{c^2} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \end{cases} \tag{79}$$

we clearly see that $x, y, z \neq 0$ and $\lambda \neq 0$. We can multiply the first three equations to obtain

$$64x^2y^2z^2 = \frac{\lambda^3 xyz}{a^2b^2c^2} \tag{80}$$

in which case

$$\lambda = 4\sqrt[3]{a^2b^2c^2xyz} \tag{81}$$

Dividing the first equation by the second equation we obtain

$$\frac{y}{x} = \frac{b^2x}{a^2y} \tag{82}$$

Since $x, y, z > 0$ we conclude that

$$y = \frac{b}{a}x \tag{83}$$

Similarly, we divide the first equation by the third one to obtain

$$z = \frac{c}{a}x \tag{84}$$

Substituting this information in the equation for the ellipsoid we find that

$$3\frac{x^2}{a^2} = 1 \tag{85}$$

Which gives us the values

$$x = \frac{a}{\sqrt{3}} \quad y = \frac{b}{\sqrt{3}} \quad z = \frac{c}{\sqrt{3}} \tag{86}$$

and from this we conclude that the volume must be $\frac{8abc}{3\sqrt{3}}$.

Example 9. Find the absolute maxima and minima of the function $f(x, y, z) = x+y+z$ inside the region $A = \left\{(x, y, z) : \ x^2 + y^2 + z^2 \leq 1\right\}$

Since this set is closed and bounded then an absolute maximum and minimum will be achieved. First we work with the interior of the sphere, that is, the region $x^2+y^2+z^2 < 1$. Here it is easy to see that $\nabla f(x, y, z) = \mathbf{0}$ has no solutions.

Therefore we can focus on the region $x^2+y^2+z^2 = 1$. In this case we use the constrain $g(x, y, z) = x^2 + y^2 + z^2 - 1$ and so the method of Lagrange multipliers requires us to solve $\nabla f = \lambda \nabla g$, that is

$$(1, 1, 1) = \lambda\,(2x, 2y, 2z) \tag{87}$$

We obtain the system of equations

$$\begin{cases} 1 = 2\lambda x \\ 1 = 2\lambda y \\ 1 = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases} \tag{88}$$

Again, $\lambda \neq 0$ so $x = y = z = \frac{1}{2\lambda}$. Substituting in the last equation we find $x = \pm\frac{1}{\sqrt{3}}$ which means that the critical points are $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ and $\left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$.

Since the maximum and minimum must be achieved and there are only two candidates $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ will correspond to the absolute maximum while $\left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$ will correspond to the absolute minimum.

Problem 10. Find the dimensions of a box (top included) which contains a given volume V and uses minimum material (i.e, has minimum surface area) using the Lagrange multipliers method.

Notice that we solved this problem before, but now we are going to use the Lagrange multipliers method. The function we want to minimize is the surface area [called before S]

$$f(x, y, z) = 2xy + 2xz + 2yz \tag{89}$$

subject to the constraint equation

$$g(x, y, z) = xyz - V = 0 \tag{90}$$

According to the Lagrange multiplier method, we must solve

$$\begin{cases} \nabla f = \lambda \nabla g \\ g = 0 \end{cases} \tag{91}$$

which is equivalent to

$$\begin{cases} (2y + 2z, 2x + 2z, 2x + 2y) = \lambda(yz, xz, xy) \\ xyz = V \end{cases} \tag{92}$$

We obtain the system of equations

$$\begin{cases} 2y + 2z = \lambda yz & (1) \\ 2x + 2z = \lambda xz & (2) \\ 2x + 2y = \lambda xy & (3) \\ xyz = V & (4) \end{cases} \tag{93}$$

To solve this system, multiply equation (1) by x, equation (2) by y , and equation (3) by z, in order to obtain

$$\begin{cases} 2xy + 2xz = \lambda xyz & (A) \\ 2xy + 2yz = \lambda xyz & (B) \\ 2xz + 2yz = \lambda xyz & (C) \\ xyz = V & (4) \end{cases} \tag{94}$$

Now use the last equation and substitute in the other three to obtain

$$\begin{cases} 2xy + 2xz = \lambda V & (A) \\ 2xy + 2yz = \lambda V & (B) \\ 2xz + 2yz = \lambda V & (C) \end{cases} \tag{95}$$

The left hand side of (A) must now equal the left hand side of (B), which means

$$\begin{aligned} & 2xy + 2xz = 2xy + 2yz \\ \Longrightarrow\ & 2xz = 2yz \\ \Longrightarrow\ & \boxed{x = y} \end{aligned}$$

Notice that here we do not need to worry about the case where $z = 0$, since that is not physically interesting. Likewise, equating (A) with (C) you will find that

$$\boxed{y = z} \tag{96}$$

So we conclude that

$$\boxed{x = y = z} \tag{97}$$

We can substitute back this information in equation (4) to find that

$$x^3 = V \tag{98}$$

or

$$\boxed{x = y = z = V^{1/3}} \tag{99}$$

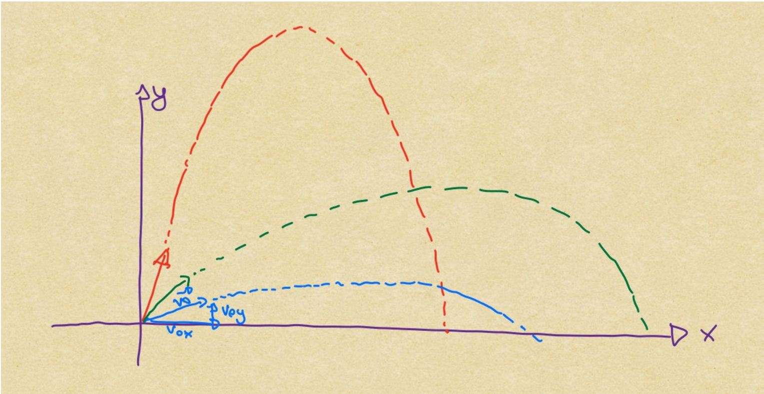
which is what we found before! The main differences to observe are:

- We treated all the equations in a symmetric fashion. That is, we did not eliminate one of the variables in terms of the others, as we had done originally when we wrote $z = \frac{V}{xy}$.
- In this particular problem we did not need to find the value of λ before obtaining the values for x, y, z. Sometimes it will be necessary to find λ first before being able to completely solve the problem.
- We found a critical point, but strictly speaking we do not know if this yields a local max, local min, or neither. It is possible to classify critical points using the Lagrange multipliers method, but we will avoid doing this since it is more complicated.

Physical Interpretation of the Lagrange Multipliers [optional]

So far it is not clear if the Lagrange multiplier λ has any particular meaning. To discuss the physical interpretation of the Lagrange multiplier, consider the following experiment.

Suppose you have a cannon and start launching balls from it with different initial velocities $\mathbf{v}_0$. Depending on the initial velocity chosen, the ball will hit the floor at different distances from the cannon, in other words, the range of each projectile will depend on the initial velocity.



In fact, if you write the initial velocity in terms of its components in the x and y directions, that is,

$$\mathbf{v}_0 = (v_{0x}, v_{0y}) \tag{100}$$

Then it is a simple physics exercise (assuming there is no friction, etc) to show that the range R of the projectile is given as

$$R(v_{0x}, v_{0y}) = \frac{2}{g} v_{0x} v_{0y} \tag{101}$$

where g is the acceleration due to gravity (a constant).

Now, if we were interested in *maximizing* the range R, then clearly a constraint needs to be imposed, since otherwise you could keep launching the projectiles with arbitrarily large velocities, thereby increasing the range indefinitely. Since the total energy is conserved in this problem, and it must equal the initial kinetic energy of the balls, it is reasonable to impose the condition that we will only launch balls with a particular kinetic energy. That is, our constraint equation is

$$C(v_{0x}, v_{0y}) = \frac{1}{2} m (v_{0x}^2 + v_{0y}^2) - E = 0 \tag{102}$$

where E is the kinetic energy. Therefore, we have naturally found a Lagrange multiplier problem: we want to optimize R subject to the constraint $C = 0$. According to Lagrange's method, we must solve

$$\begin{cases} \nabla R = \lambda \nabla C \\ C = 0 \end{cases} \tag{103}$$

This is equivalent to

$$\begin{cases} \frac{2}{g} v_{0y} = \lambda m v_{0x} & (1) \\ \frac{2}{g} v_{0x} = \lambda m v_{0y} & (2) \\ \frac{1}{2} m (v_{0x}^2 + v_{0y}^2) = E & (3) \end{cases} \tag{104}$$

Similar to the box problem, in order to solve this system we multiply equation (1) by

v_{0y}, equation (2) by v_{0x} in order to obtain

$$\begin{cases} \frac{2}{g}v_{0y}^2 = \lambda m v_{0x}v_{0y} & (A) \\ \frac{2}{g}v_{0x}^2 = \lambda m v_{0y}v_{0x} & (B) \\ \frac{1}{2}m(v_{0x}^2 + v_{0y}^2) = E & (3) \end{cases} \tag{105}$$

We equate the left hand sides of (A) and (B) to obtain

$$\frac{2}{g}v_{0y}^2 = \frac{2}{g}v_{0x}^2 \tag{106}$$

which in this case gives

$$\boxed{v_{0y} = v_{0x}} \tag{107}$$

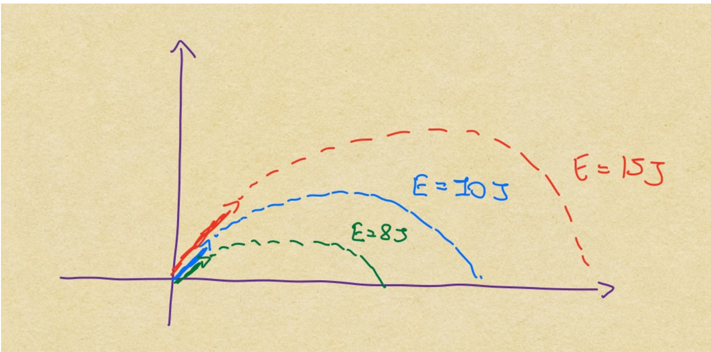
Notice that the case $v_{0y} = -v_{0x}$ is also possible but not interesting in this particular situation. So we just found that the initial velocity which maximizes the range is the one where the ball is launched at a 45° angle. Using equation (3) we can determine v_{0x} in terms of E as

$$\boxed{v_{0x} = v_{0y} = \sqrt{\frac{E}{m}}} \tag{108}$$

Again, in this case we did not find the value of λ at all, but using equation (A) and the fact that $v_{0x} = v_{0y}$ we can see that

$$\boxed{\lambda = \frac{2}{mg}} \tag{109}$$

In order to interpret this particular expression for λ, we must do a slightly different experiment.



Now that we know that for a given value of E, the velocity that maximizes the range R is one where $v_{0x} = v_{0y}$, we can do an experiment where we launch projectiles with *different* values of E, but all being launched at an 45° angle, which is the angle which maximizes the range.

Since we are fixing the angle, the range now depends solely on the energy E chosen for each launch. In fact, the formulas $v_{0x} = v_{0y} = \sqrt{\frac{E}{m}}$ and $R(v_{0x}, v_{0y}) = \frac{2}{g}v_{0x}v_{0y}$ we can think of R as given in terms of E:

$$R(E) = \frac{2}{g}\frac{E}{m} \tag{110}$$

In particular, notice that

$$\frac{dR}{dE} = \frac{2}{gm} = \lambda \tag{111}$$

This is not a coincide, in fact, we can think of the Lagrange multiplier in the following way:

Physical Meaning of the Lagrange Multiplier:
The Lagrange multiplier is the rate of change of the optimized quantity R with respect to value of the constraint parameter E.

Lagrange Multipliers and Absolute Max/Min Problems

Suppose you want to find the absolute maxima and minima of a function f on some bounded region R. Moreover, suppose that the boundary of the region R is given by some equation $g = 0$, which is interpreted as the constraint equation.

1. Find the critical points of the function f, that is, the points where $\nabla f = \mathbf{0}$. Make a list with the critical points which belong to the region R.

2. Find the critical points given by the Lagrange multiplier problem, that is, solve $\nabla f = \lambda \nabla g$.

3. Make a table with the critical points from steps 1. and 2., and evaluate f at all of these points. The absolute max and min will correspond the largest and smallest values of f that appear on this table.

Example 11. Find the absolute maxima and minima of the function $f(x, y) = 8x^2 - 2y^2$ when restricted to the region R on the xy plane given by the inequality $x^2 + y^2 \leq 1$.
We follow the steps:

1. The gradient of f is $\nabla f = (16x, -4y)$ so $\nabla f = \mathbf{0}$ only when $x = 0$ and $y = 0$. Therefore, $(0, 0)$ is the only (ordinary) critical point. Moreover, notice that $(0, 0)$ belongs to the region R so we will keep it.

2. We will solve $\nabla f = \lambda \nabla g$, where $g(x, y) = x^2 + y^2 - 1$ is the equation of the circle. Since $\nabla g = (2x, 2y)$, we must solve

$$\begin{cases} 16x = 2\lambda x \\ -4y = 2\lambda y \\ x^2 + y^2 = 1 \end{cases} \tag{112}$$

The first equation is the same as $2x(8 - \lambda) = 0$. **Case** $x = 0$**:** then the third equation implies that $y = \pm 1$ and so the critical points are $(0, 1)$ and $(0, -1)$. **Case** $\lambda = 8$**:** the second equation becomes $-4y = 16y$ so $y = 0$, and the third equation implies that $x = \pm 1$. Therefore, the critical points are $(1, 0)$ and $(-1, 0)$.

3. The table with all the critical points is

$$\begin{array}{cc} (x,y) & f(x,y) = 8x^2 - 2y^2 \\ & \\ (0,0) & 0 \\ (0,1) & -2 \\ (0,-1) & -2 \\ (1,0) & 8 \\ (-1,0) & 8 \end{array} \tag{113}$$

So $(0,1)$ and $(0,-1)$ yield the absolute minimum, which is -2, while $(1,0)$ and $(-1,0)$ yield the absolute maximum, which is 8. Notice that in this case it is not necessary to apply the second derivative test to $(0,0)$, since that would only say whether or not it yields a *relative* max, *relative* min, or a saddle point, which is different from being an absolute max or min.