polar coordinates:

Every point $(x, y) \in R^2$ other than $(0, 0)$ can be uniquely expressed as $x = r\cos\theta; y = r\sin\theta$ for some (r, θ) with $0 < r < \infty$ and $0 \leq \theta < 2\pi$. these (r, θ) are called polar coordinates of the cartesian point (x, y). The Jacobian of the transformation is r. Thus when you integrate $dxdy$ is transformed to $rdrd\theta$.

Every point $(x, y, z) \in R^3$ other than $(0, 0, 0)$ can be uniquely expressed as

$$x = r\cos\theta; \quad y = r\sin\theta\cos\psi; \quad z = r\sin\theta\sin\psi.$$

where

$$0 < r < \infty; \quad 0 \leq \theta \leq \pi; \quad 0 \leq \psi < 2\pi.$$

These (r, θ, ψ) are called the spherical or polar coordinates of (x, y, z). The Jacibian is $r^2 \sin\theta$. Thus when you integrate

$$f(x, y, z)\ dx\ dy\ dz$$

is transformed to

$$f(r, \theta, \psi)\ r^2\ \sin\theta\ dr\ d\theta\ d\psi.$$

In n dimensions lo there is a similar transformation to polar coordinates. Every point

$$(x_1, x_2, \cdots, x_n)$$

can be uniquely expressed as

$$x_1 = r\cos\theta_1;$$

$$x_2 = r\sin\theta_1\cos\theta_2;$$

$$x_3 = r\sin\theta_1\sin\theta_2\cos\theta_3;$$

$$\vdots$$

$$x_{n-1} = r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\cos\theta_{n-1};$$

$$x_n = r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\sin\theta_{n-1};$$

where

$$0 < r < \infty;\ \theta_1, \theta_2, \cdots \theta_{n-2} \in [0, \pi];\ \theta_{n-1} \in [0, 2\pi).$$

The Jacobian of this transformation is

$$r^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}$$

The proof is by induction on n. You can write the Jacobian and expand as sum of two determinants for each of which the induction hypothesis applies.

Thus

$$dx_1dx_2\cdots dx_n$$

is transformed to

$$r^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}\;drd\theta_1\cdots d\theta_{n-1}.$$

Here the sphere

$$\{x : 0 < ||x|| \leq R\}$$

is transformed to the rectangle

$$[0,R]\times[0,\pi]\times\cdots\times[0,\pi]\times[0,2\pi).$$

The representation itself is also proved by induction. You need to put

$$r = \sqrt{\sum x_i^2}.$$

you need only represent the unit vector $(u_i) = (x_i/r)$. You get a θ_1 such that

$$u_1 = \cos\theta_1; \quad \sqrt{\sum_2^n u_i^2} = \sin\theta_1.$$

Since both these numbers are non-negative a number $\theta_1 \in [0,\pi]$ exists and it is unique. Now use induction hypothesis.

cylindrical coordinates:

Given a point $(x,y,z) \in R^3$ such that $(x,y) \neq (0,0)$ there is a unique number $0 < r < \infty$ and $0 \leq \theta < 2\pi$ such that $x = r\cos\theta; y = r\sin\theta$. the triple (r,θ,z) are called the cylindrical coordinates of the point (x,y,z). jacobian of this transformation is r. Thus

$$dxdydz$$

is transformed to

$$rdrd\theta dz.$$

this is easy to establish. Here the cylinder (z-axis deleted)

$$\{(x, y, z) : 0 < x^2 + y^2 \leq R^2; 0 \leq z \leq h\}$$

is transformed to the rectangle

$$\{(r, \theta, z) : 0 < r \leq R; 0 \leq \theta <, 2\pi; 0 \leq z \leq h\}.$$

Lagrange multipliers:

We have discussed analogue of maxima and minima in two variables. Many a times you need to optimize (find local extrema of) a function subject to certain constraints. For example you want to find out in the positive quadrant

$$\max(x + y) \quad \text{subject to} \quad x + y = 1$$

(actually you need not say positive quadrant, it follows even if you did not say it).

Of course in the above problem, you can say $y = 1 - x$ and we need to maximize $x(1 - x)$. Manytimes such a simplification is not possible. Rewrite the above problem as

$$\max f(x, y) \quad \text{subject to} \quad \varphi(x, y) = 0$$

where $f(x, y) = xy$ and $\varphi(x, y) = x + y - 1$. Lagrange found out that at such a point (x_0, y_0) where there is a max (or min) there is a number λ such that

$$\nabla f + \lambda \nabla \varphi = 0.$$

This method works even in complicated situations when you can not explicitly express one variable as a function of the other variable. More over the above equations maintain a symmetry without expressing one variable as a function of the other. Thus you can solve the three equations, namely, above system of two equations along with the one equation $\varphi = 0$ to obtain the three numbers (x_0, y_0, λ).

Of course you would ask what is the worth of this? You need to find out precisely the points of max and min and Lagrange only tells us what happens if you already know that (x_0, y_0) is such a point. You can utilise the above discovery to first solve the system of equations and get all solutions

(x, y, λ) and reduce your search to check only among these points. Usually, the system does not have too many solutions and so it is easy to check this.

Of course, you need to check whether a solution you obtained is max or min or neither. There is no simple criterion. remember Lagrange only tells you at a point of (constrained) local extremum something happens. he does not say that if such a thing happens then the point is an extremum. We see examples.

to proceed further, let us define what is meant by constrained extremum. Let f and φ be two real valued functions on an open set $\Omega \subset R^2$. A point $P \in \Omega$ is a local maximum of f subject to the constraint $\varphi = 0$ if there is an $\epsilon > 0$ such that the following happens:

$$\varphi(P) = 0;$$

and

$$Q \in \Omega;\ \varphi(Q) = 0;\ ||P - Q|| < \epsilon\ \Rightarrow f(Q) \leq f(P).$$

Thus locally at P there is no other point satisfying the constraint and gives a larger value for f.

Similarly we define local minimum subject to the constraint. These are called local extrema subject to the given constraint.

Theorem: f and φ be C^1 functions on an open set $\Omega \subset R^2$; $P = (x_0, y_0) \in \Omega$ is a constrained local extremum. Assume that $\varphi_y(P) \neq 0$.

Then there is a number λ such that

$$\nabla f(P) + \lambda \nabla \varphi(P) = 0.$$

Proof is simple. Assume we have a local max.

Since $\varphi(P) = 0$ and $\varphi_y(P) \neq 0$, the implicit function theorem applies. We have a rectangle $(a, b) \times (c, d)$ which includes P and has the following property. For every $x \in (a, b)$ there is a unique $y \in (c, d)$ such that $\varphi(x, y) = 0$. If you define $g(x)$ as this unique y, then g defines a C^1 function on (a, b). In other words the set of points $(x, g(x))$ captures *all* zeros of the φ in this rectangle.

If necessary we can take a smaller rectangle so that $\varphi(P)$ is max in this rectangle. Since g captures all zeros of φ in the sense described above we conclude that the function $f(x, g(x))$ assumes its max value on the interval (a, b) at the point x_0 and hence its derivative is zero at this point. Thus chain rule for the map

$$x \quad \mapsto \quad (x, g(x)) \quad \mapsto \quad f(x, g(x)).$$

We get

$$f_1(P) + g'(x_0) f_2(P) = 0. \qquad (\bullet)$$

How do we get rid of g that we introduced from the above. apply chain rule to the map

$$x \quad \mapsto \quad (x, g(x)) \quad \mapsto \varphi(x, g(x)) \equiv 0.$$

$$\varphi_x(P) + g'(x_0) \varphi_y(P) = 0. \qquad (*)$$

(You can apply the formula for derivative of function defined implicitly). Combining $(\bullet)$ and $(*)$

$$\begin{pmatrix} f_1(P) & f_2(P) \\ \varphi_x(P) & \varphi_y(P) \end{pmatrix} \begin{pmatrix} 1 \\ g'(x_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Thus we have a matrix equation $Av = 0$ where the vector v is non-zero. Thus the rows of the matrix must be linearly dependent. Since we know $\varphi_y(P) \neq 0$, the second row is non-zero row. So the first row must be a multiple of the second row. The constant denoted by $(-\lambda)$ satisfies the requirement.

This completes the proof.

Let us see some examples.

(1) maximize $f(x, y) = xy$ subject to $x^2 + y^2 = 1$.

Here $\varphi(x, y) = x^2 + y^2 - 1$. The equations reduce to

$$y + 2\lambda x = 0; \quad x + 2\lambda y = 0; \quad x^2 + y^2 = 1.$$

Multiply first eqn by x, second by y; add; use third to see $xy + \lambda = 0$. Use this in the first. You will see the only solutions are

$$(x, y, \lambda) = (1, 0, 0); (0, 1, 0); (1/\sqrt{2}, 1/\sqrt{2}, 1/2).$$

The first two are not extrema. third is.

(2) Maximize $f(x, y, z) = xyz$ subject to $x^2 + y^2 + z^2 = 1$.

Exactly the same procedure leads to

$$yz + 2\lambda x = 0; \quad xy + 2\lambda z = 0; \quad zx + 2\lambda y = 0; \quad x^2 + y^2 + z^2 = 1$$

so that Multiply first eqn by x etc and add; use the constraint to see

$$2\lambda = -3xyz.$$

substitute in the above three eqns to get the solutions

$$(\pm 1, 0, 0) \ \ (0, \pm 1, 0) \ \ (0, 0, \pm 1) \quad \left(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right).$$

It is easy to see that the first six solutions are neither max nor min. At these points f is zero. In the neighbourhood of these points f takes values positive and also negative. Of the next four points four are maxima and four minima.

In an extremal problem where $\nabla f(P) = 0$, if there are points Q in every neighbourhood of P such that $f(P) > f(Q)$ as well as there are points Q in every neighbourhood of P with $f(P) < f(Q)$; then we say p is a saddle point.

Similarly in a constrained extremal problem suppose we have a solution (P, λ) for the Lagrange method, that is, satisfying $\nabla f(P) + \lambda\varphi(P) = 0$. Suppose that in every neighbourhood of P there are points Q satisfying the constraint and $f(P) > f(Q)$ as well as points Q satisfying the constraint such that $f(P) < f(Q)$. Then we say that P is saddle point for the constrained problem.

In the above problem, the first six points are saddle points.

(3) We are given an $n \times n$ symmetric matrix A.

maximize $x^t Ax$ subject to $||x||^2 = 1$.

Note that the set $||x|| = 1$ is a compact set and and the expression has a maximum.

The expression $f(x) = x^t Ax = \sum_{i,j} a_{ij}x_i x_j$ is called a quadratic form. Here $\varphi(x) = \sum x_i^2 - 1$.

The equations are, denoting the lagrange constant by $(-\lambda)$,

$$2\sum_j a_{ij}x_j - 2\lambda x_i = 0; \quad i = 1, 2, \cdots, n.$$

That is

$$Ax - \lambda x = 0; \quad Ax = \lambda x.$$

In other words the Lagrange constant λ is an eigen value and x corresponding eigen vector. But which eigen value is it? At this point

$$x^t A x = x^t \lambda x = \lambda.$$

Thus Lagrange method helps you to identify the largest eigen value as maximum of the quadratic form. Of course, there are several eigen values and corresponding eigen vectors. We know a priori that there is a maximum and hence this procedure gives it along with others.

(4) Find box with sides parallel to the coordinate planes which has the largest volume and is contained in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1.$$

A box (or rectangular parallelopiped) is a figure with rectangular sides. It may not be having sides parallel to the coordinate planes. For example start with a box with sides parallel to the coordinate planes and apply a rotation. What you get is still a box with rectangular faces. It should be possible to show that such a box contained in the above ellipsoid, having largest volume must actually have sides parallel to the coordinate planes. I do not have a proof right now.

Returning to our problem let us consider boxes as stated. If none of the corners is on the boundary of the ellipsoid, then you can enlarge the box increasing the volume. If one corner (x_0, y_0, z_0) is on the boundary then the box must have corners $(\pm x_0, \pm y_0, \pm z_0)$ and hence volume $8x_0y_0z_0$. Thus the geometric problem can be formulated as the following analytical problem.

Maximize $8xyz$ subject to $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Proceeding exactly as above you get corners to be $(a/\sqrt{3}, b/\sqrt{3}, c/\sqrt{3})$. and hence the volume is $8abc/(3\sqrt{3})$.

Lagrange with two constraints:

Lagrange method works with any variables with reasonable number of constraints. I shall state the general theorem. But before that I shall state for three variables with two constraints.

Maximize $f(x, y, z)$ subject to $\varphi(x, y, z) = 0$; $\psi(x, y, z) = 0$.

We assume all are C^1 functions defined on an open set Ω. We assume that $P = (x_0, y_0, z_0) \in \Omega$ is a solution for the problem. We show there are numbers λ and μ such that

$$\nabla f(P) + \lambda \nabla \varphi(P) + \mu \nabla \psi(P) = 0.$$

Of course, this comes with a price. We assume that

$$\begin{pmatrix} \varphi_y(P) & \varphi_z(P) \\ \psi_y(P) & \psi_z(P) \end{pmatrix} \quad (\spadesuit)$$

is non-singular, that is, the determinant is non zero.

Under the hypothesis, we can apply implicit function theorem. get an interval $I = (x_0 - \delta, x_0 + \delta)$ and an box (rspa) Q such that

(i) $(y_0, z_0) \in Q$; $\quad I \times Q \subset \Omega$.

(ii) for $x \in I$ there is unique $(y, z) \in Q$ such that $\varphi(x, y, z) = 0$ and $\psi(x, y, z) = 0$. Moreover the functions $g : I \to Q$ that maps $x \mapsto (y, z)$ is a C^1 map. Denote $g(x) = (g_1(x), g_2(x))$. Thus $g(x_0) = (y_0, z_0)$.

Thus region $I \times Q$ contains (x_0, y_0, z_0) and all common zeros of φ and ψ are captured by $\{(x, g(x)) : x \in I\}$ in this region. In other words x_0 is a extremal point for the function $f(x, g(x))$. so derivative of this function must be zero at x_0. applying chain rule

$$x \mapsto (x, g_1(x), g_2(x)) \mapsto f(x, g_1(x), g_2(x))$$

$$f_x(P) + f_y(P) g_1'(x_0) + f_z(P) g_2'(x_0) = 0. \quad (*)$$

Applying chain rule to

$$x \mapsto \langle x, g_1(x), g_2(x) \rangle$$

$$\mapsto \langle \varphi(x, g_1(x), g_2(x)), \psi(x, g_1(x), g_2(x)) \rangle \equiv \langle 0, 0 \rangle.$$

$$\begin{pmatrix} \varphi_x(P) & \varphi_y(P) & \varphi_z(P) \\ \psi_x(P) & \psi_y(P) & \psi_z(P) \end{pmatrix} \begin{pmatrix} 1 \\ g_1'(x_0) \\ g_2'(x_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (**)$$

combining $(*)$ and $(**)$ we see

$$\begin{pmatrix} f_x(P) & f_y(P) & f_z(P) \\ \varphi_x(P) & \varphi_y(P) & \varphi_z(P) \\ \psi_x(P) & \psi_y(P) & \psi_z(P) \end{pmatrix} \begin{pmatrix} 1 \\ g_1'(x_0) \\ g_2'(x_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We have a matrix equation $Av = 0$ with a non-zero vector v. so A has dependent rows. Last two rows are independent, because y hypothesis, the 2×2 bottom corner is non-singular. Thus the first row must be a linear combination of the last two. This gives λ and μ.

Instead of (♠) we could have assumed

$$\begin{pmatrix} \varphi_z(P) & \varphi_x(P) \\ \psi_z(P) & \psi_x(P) \end{pmatrix}$$

is non-singular. Then we get an interval around y_0 etc. the same proof works. Or we could assume, instead,

$$\begin{pmatrix} \varphi_x(P) & \varphi_y(P) \\ \psi_x(P) & \psi_y(P) \end{pmatrix}$$

is non-singular.

general Lagrange:

Let $f, \varphi_1, \cdots \varphi_m$ be all real valued C^1 functions defined on an open set $\Omega \subset R^n$. Assume $m < n$. Suppose $x^0 \in \Omega$ is an extremal for f subject to

$$\varphi_i(x) = 0; \quad i = 1, 2, \cdots, m.$$

Then there are numbers $(\lambda_i : 1 \leq i \leq m)$ such that

$$\nabla f(x^0) + \lambda_1 \nabla \varphi_1(x^0) + \lambda_2 \nabla \varphi_2(x^0) + \cdots + \lambda_m \nabla \varphi_m(x^0) = 0.$$

This comes with a price. We need to assume that the matrix

$$\left(\frac{\partial}{\partial x_j} \varphi_i(x^0) : 1 \leq i, j \leq m \right)$$

is non-singular. we have considered the $m \times m$ matrix by differentiating the constraining function w.r.t. the first m coordinates. You can take any m coordinates and use them for all the functions. Demand that this matrix be non-singular. Remember we are evaluating the matrix at the point x^0.

you should keep in mind two things. the method reduces the search for extremals among solutions of the following $n + m$ equations.

$$\nabla f(x) + \lambda_1 \nabla \varphi_1(x) + \lambda_2 \nabla \varphi_2(x) + \cdots + \lambda_m \nabla \varphi_m(x) = 0.$$

$$\varphi(x) = 0; \varphi_2(x) = 0; \cdots \varphi_m(x) = 0.$$

solve these $n+m$ equations to get solutions

$$(x, \lambda) : x \in R^n; \lambda \in R^m$$

and search those x for max or min. of course not all may be extremals. But extremals are definitely contained among these.