Unit 9: Hyperplanes

Juan Luis Melero ans Eduardo Eyras

October 2018
1 Hyperplanes

1.1 Definition

A hyperplane in an $n$–dimensional vector space $\mathbb{R}^n$ is defined to be the set of vectors:

$$u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

satisfying the equation:

$$a_1 x_1 + \cdots + a_n x_n = b$$

where $a_1, \ldots, a_n$ and $b$ are real numbers with at least $a_1, \ldots, a_n$ non-zero.

For example, in $\mathbb{R}^2$ a hyperplane is a line:

\[ x + y = 4 \]

Figure 1: Graphical representation of the hyperplane equation $x + y = 4$

And for example, in $\mathbb{R}^3$, a hyperplane is a 2D plane:
1.2 Hyperplane equations

To define the hyperplane equation we need either a point in the plane and a unit vector orthogonal to the plane, two vectors lying on the plane or three coplanar points (they are contained in the hyperplane). Is is usually built from the combination of a point and a vector and it corresponds to a $n - 1$ dimensional vector subspace.

For the first case, we will get the equation from a point and an orthogonal unit vector.

The vector $p - p_0$ lies on the plane, thus it is orthogonal to $n$:

$$n \perp (p - p_0) \implies \langle n, p - p_0 \rangle = 0$$

This provides the vector equation of the plane: $\langle n, p - p_0 \rangle = 0$. Given the vector:

$$n = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$
Performing the scalar product and taking common factors, we end up with the equation:
\[ a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \]
This is the scalar equation of the plane. It is also possible to write it in terms of a constant value \( D \):
\[ D = ax_0 + by_0 + z_0 \implies ax + by + cz = D \]
This is the general equation of the plane.

As \( n \) defines the normal line to the hyperplane, we can define a normal line with the normal vector:
\[ r(t) = p_0 + tn \iff \begin{cases} x(t) = x_0 + at \\ y(t) = y_0 + bt \\ z(t) = z_0 + ct \end{cases} \]
for \( n = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \).
Figure 4: Graphical representation of a line, which is a $\mathbb{R}^2$ hyperplane. $v$ represents the vector in the direction of the hyperplane. $p$ represents the point which the hyperplane goes through.

**Example:** line in $\mathbb{R}^2$ in the direction of $\vec{v} = (1, 1)$ and going through the point $p = (1, 0)$.

Any point in the line is given by the equation:

$$r = p + \alpha v = \begin{cases} x = p_1 + \alpha v_1 = 1 + \alpha \\ y = p_2 + \alpha v_2 = 0 + \alpha \end{cases}$$

This is the parametric equation. We can isolate the parameter $\alpha$ and have the equation:

$$r = p + \alpha v = \begin{cases} x = p_1 + \alpha v_1 = 1 + \alpha \\ y = p_2 + \alpha v_2 = 0 + \alpha \end{cases} \rightarrow y = x - 1$$

This is the intrinsic equation.

We can build it as a hyperplane. Calculate a vector orthogonal to the line:

$$\langle n, v \rangle = u_1 + u_2 = 0 \rightarrow n = \begin{pmatrix} u_1 \\ -u_1 \end{pmatrix} \rightarrow n = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \|n\| = 1$$

So, with the orthogonal vector and the specific known point $p$, we can build the equation for any point $r$ of the hyperplane:

$$\langle n, r \rangle = \langle n, p \rangle$$
\[
\begin{pmatrix}
\frac{1}{\sqrt{2}} & -1/\sqrt{2}
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
= \begin{pmatrix}
\frac{1}{\sqrt{2}} & -1/\sqrt{2}
\end{pmatrix}
\begin{pmatrix}
1 \\
0
\end{pmatrix}
\rightarrow
\]
\[
\frac{1}{\sqrt{2}}(x - y) = \frac{1}{\sqrt{2}} = y = x - 1 \text{ the same as before}
\]

If we want to build the normal line through the point, we use the normal vector and the parametric equation:

\[
n = \begin{pmatrix}
\frac{1}{\sqrt{2}} \\
-1/\sqrt{2}
\end{pmatrix}
\rightarrow r(t) = p_0 + tn \rightarrow \begin{cases}
x(t) = x_0 + t \frac{1}{\sqrt{2}} \\
y(t) = y_0 - t \frac{1}{\sqrt{2}}
\end{cases}
\]

Using the same point \( p_0 = (1, 0) \) as before, we get:

\[
\begin{cases}
x(t) = 1 + t \frac{1}{\sqrt{2}} \\
y(t) = -t \frac{1}{\sqrt{2}}
\end{cases}
\rightarrow y = 1 - x
\]

Figure 5: Graphical representation of the normal line of an \( \mathbb{R}^2 \) hyperplane. \( v \) represents the vector in the direction of the hyperplane. \( p \) represents the point which the hyperplane and the normal line go through. \( y = 1 - x \) is the equation of the normal line.

**Example in \( \mathbb{R}^3 \):** Consider the normal vector \( \vec{n} \) and the vector \( p_0 \) representing a given point in the plane:

\[
\vec{n} = \begin{pmatrix}
4 \\
-1 \\
6
\end{pmatrix}, \quad p_0 = \begin{pmatrix}
0 \\
1 \\
-7
\end{pmatrix}
\]
We can write the equation of the hyperplane defined by these two vectors:

\[ \langle n, p - p_0 \rangle = 0 \rightarrow \]
\[ a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \rightarrow \]
\[ 4(x - 0) - (y - 1) + 6(z + 7) = 0 \rightarrow \]
\[ 4x - y + 6z = -43 \]

where \( p \) represents a vector for a generic point in the plane:

\[ p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \]

As we said before:

\[ a(x-x_0)+b(y-y_0)+c(z-z_0) = 0 \iff ax+by+cz = D, \text{ with } D = ax_0+by_0+cz_0 \]

There is no point \((x, y, z)\) can fulfill two conditions:

\[ ax + by + cz = D \]
\[ ax + by + cz = D' \]
\[ D \neq D' \]

So there are no common points, i.e. no intersection \(\rightarrow\) they are parallel.

A hyperplane in \( \mathbb{R}^n \) can be represented by a single equation:

\[ a_1x_1 + a_2x_2 + \cdots + a_nx_n = D \]

where \( n = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \) is a vector orthogonal to the hyperplane as we chose a point \( p_0 \) in the plane such that \( \langle n, p_0 \rangle = D \).

All planes of that similar form, for any value of \( D \) are parallel to each other. In particular, for \( D = 0 \) we have \( a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \), which can be re-written in temps of a \( 1 \times n \) matrix \( A^T = n \):

\[ \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \rightarrow Au = 0 \]
Hence, the hyperplane is the kernel of the matrix $A$: $Ker(A)$. The hyperplane is a vector subspace. In addition, considering that $\text{rank}(A) = 1$ and $\dim(\mathbb{R}^n) = n$ then we see that the hyperplane has dimension $n - 1$:

$$\text{rank}(A) + \dim(Ker(A)) = n \rightarrow \dim(Ker(A)) = n - 1$$

### 1.3 Vector product

Vector product or the cross product is an operation that allows to find a vector perpendicular to two known vectors. This is useful here because to build the hyperplane equation we may need it.

Given two vectors in $\mathbb{R}^3$, $u, v \in \mathbb{R}^3$ and the canonical basis $E_3 = \{e_1, e_2, e_3\}$, we define the vector product of these two vectors as:

$$u \times v = \begin{vmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (\text{we solve it with the adjoint method})$$

$$= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} e_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} e_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} e_3$$

Properties:

1. $u \times v$ is orthogonal to both vectors
2. If $u$ and $v$ are linearly dependent, then $u \times v = 0$
3. $u \times v = -v \times u$

Proof of property 1:

$$\langle u, u \times v \rangle = \langle u_1 e_1, (u_2 v_3 - u_3 v_2) e_1 \rangle + \langle u_2 e_2, (u_3 v_1 - u_1 v_3) e_2 \rangle + \langle u_3 e_3, (u_1 v_2 - u_2 v_1) e_3 \rangle =$$

$$= u_1(u_2 v_3 - u_3 v_2) \langle e_1, e_1 \rangle + u_2(u_3 v_1 - u_1 v_3) \langle e_2, e_2 \rangle + u_3(u_1 v_2 - u_2 v_1) \langle e_3, e_3 \rangle =$$

$$= u_1(u_2 v_3 - u_3 v_2) + u_2(u_3 v_1 - u_1 v_3) + u_3(u_1 v_2 - u_2 v_1) =$$

$$= u_1 u_2 v_3 - u_1 u_3 v_2 + u_2 u_3 v_1 - u_2 u_1 v_3 + u_3 u_1 v_2 - u_3 u_2 v_1 = 0$$

Similar for $\langle v, u \times v \rangle$. 
Proof of property 2: (recall the properties of determinants)

\[ u \times v = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ \alpha u_1 & \alpha u_2 & \alpha u_3 \end{pmatrix} = 0 \]

Proof of property 3: (recall the properties of determinants)

\[ u \times v = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} = -\det \begin{pmatrix} e_1 & e_2 & e_3 \\ v_1 & v_2 & v_3 \\ u_1 & u_2 & u_3 \end{pmatrix} = -v \times u \]

Example: determine the equation of the plane in \( \mathbb{R}^3 \) that contains the points:

\[ P = (1, -2, 0) = OP \]
\[ Q = (3, 1, 4) = OQ \]
\[ R = (0, -1, 2) = OR \]

First, we calculate two vectors in the plane:

\[ PQ = OQ - OP = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \]
\[ PR = OR - OP = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \]

Now we use the vector product of these two vectors to calculate an orthogonal vector:

\[ n = PQ \times PR = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ 2 & 3 & 4 \\ -1 & 1 & 2 \end{pmatrix} \]

\[ = \det \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} - \det \begin{pmatrix} 2 & 4 \\ -1 & 2 \end{pmatrix} + \det \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix} = 2e_1 - 8e_2 + 5e_3 \]

So the vector \( n = \begin{pmatrix} 2 \\ -8 \\ 5 \end{pmatrix} \) is orthogonal to the plane. It can be verified with the scalar product:

\[ \langle n, PQ \rangle = \begin{pmatrix} 2 & -8 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 0 \]
The same for PR.

Now we can write the equation of the plane:

\[
\langle n, r \rangle = \langle n, p \rangle
\]

\[
\langle n, p - p_0 \rangle = 0 \rightarrow a(x - x_0) + b(y - y_0) + c(z - z_0) = 0
\]

Here we choose the normal vector \( n \) calculated before and any of the three initial points, since they all are contained in the plane. We will choose the point \( P = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \), but any of the points would have been correct. So we get the equation:

\[
2(x - 1) - 8(y + 2) + 5(z - 0) = 0 \rightarrow 2x - 8y + 5z = 18
\]

2 Tangent hyperplane

2.1 Directional derivative

Partial derivatives, and overall the gradient vector, can be interpreted as the derivatives in the direction to the axis (the canonical basis vectors). For example, \( \frac{\partial f}{\partial x}(x_0, y_0) \) is the slope of the tangent line to \( f \), in the point \((x_0, y_0)\), which is parallel to the OX axis. In case of two variable functions, it is not a unique tangent line but, if it exists, is a whole tangent plane. If we want to define the derivative not in the direction of the axis but a given direction, then we calculate the directional derivative.

Directional derivative is defined as the derivative (the gradient) in the direction of a vector:

\[
D_u f = \langle \nabla f, u \rangle = u_1 \frac{\partial f}{\partial x_1} + \cdots + u_n \frac{\partial f}{\partial x_n}
\]

\( u \) is generally taken to be a unit vector.

Example: consider the function \( f(x, y) = xy \) and the vector \( u = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \).

Calculate the directional derivative along \( u \).

\[
\nabla f = \begin{pmatrix} y \\ x \end{pmatrix}
\]
\[ Du f = \langle \nabla f, u \rangle = u_1 \frac{\partial f}{\partial x_1} + \cdots + u_2 \frac{\partial f}{\partial x_2} = y + x \]

2.2 Theorems related to the tangent hyperplane

2.2.1 Maximum value of a directional derivative

Theorem: the maximum value, i.e. the maximal rate of change, of a directional derivative \( D_u f, \forall u \in \mathbb{R}^n \) is given by \( \| \nabla f \| \) and will occur in the direction of \( \nabla f \).

Proof: we use the definition of the scalar product in terms of the angle between the two vectors:

\[ D_u f = \langle \nabla f, u \rangle = \| \nabla f \| \| u \| \cos \theta = \| \nabla f \| \cos \theta \]

The maximum value of the cosine occurs at zero angle:

\[ \max_{\theta} \{ \cos \theta \} = 1 \rightarrow \theta = 0 \]

This implies that \( u \) is in the same direction as \( \nabla f \). If we maximize the directional derivative in terms of the vector \( u \):

\[ \max_u \{ D_u f \} = \max_u \{ \| \nabla f \| \cos \theta \} = \| \nabla f \| \max_{\theta} \{ \cos \theta \} = \| \nabla f \| \]

So we get that the maximum value of the directional derivative is the norm of the gradient, and it is parallel to the gradient:

\[ \max_u \{ D_u f \} = \| \nabla f \| \text{ at } u \parallel \nabla f \]

Q.E.D

2.2.2 The gradient is orthogonal to the level curve

Theorem: consider the level curve of a function \( f \), which is defined as \( f(x, y) = k \). The gradient \( \nabla f \) at a point in the level curve \( p = (x_0, y_0) \) is orthogonal to the level curve at that point \( f(x_0, y_0) = 0, k \in \mathbb{R} \). This is similar for level surfaces if we consider the function \( f(x, y, z) \) and the level surfaces \( f(x, y, z) = k, k \in \mathbb{R} \). The gradient vector of the function \( \nabla f \) at a point on the level surface \( p = (x_0, y_0, z_0) \) such that \( f(x_0, y_0, z_0) = k \) is perpendicular
to the level surface at that point.

Proof: consider the level curve of a two variables function \( f(x, y) = k \). Points in the curve can be represented by \( r(t) = (x(t), y(t)) \) (recall the parameterization of a curve). To be in the level curve, the points of the function must fulfill that \( f(x(t), y(t)) = k \). The idea is to show that the gradient at a point is orthogonal to the tangent line to the curve, which is defined as \( \frac{\partial r}{\partial t} \). First, we differentiate respect to the parameter \( t \):

\[
\frac{\partial f}{\partial t} = \frac{\partial k}{\partial t} = 0 \quad \text{k is a constant and does not depend on t}
\]

\[
\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t}
\]

(recall the chain rule)

\[
\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} \right) + \left( \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} \right) = \left( \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \right) = \left\langle \nabla f, \frac{\partial r}{\partial t} \right\rangle
\]

In summary:

\[
\frac{\partial f}{\partial t} = 0 \quad \frac{\partial f}{\partial t} = \left\langle \nabla f, \frac{\partial r}{\partial t} \right\rangle \rightarrow \left\langle \nabla f, \frac{\partial r}{\partial t} \right\rangle = 0
\]

We get that the gradient vector at a point \( p \) is orthogonal to the tangent vector to any curve that passes through \( p \) on a surface, thence the gradient is orthogonal to the surface.

We can generalize it to three variables function or even to \( n \) variable functions. We define a level surface \( f(x_1, x_2, \ldots, x_n) \). Parameterized points in the surface can be represented by \( r(t) = (x_1(t), x_2(t), \ldots, x_n(t)) \) and also must fulfill that \( f(x_1(t), x_2(t), \ldots, x_n(t)) = k \). Differentiating:

\[
\frac{\partial f}{\partial t} = 0
\]

And using the chain rule:

\[
\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial t} + \cdots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial t} = \left( \begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{array} \right) \left( \begin{array}{c} \frac{\partial x_1}{\partial t} \\ \frac{\partial x_2}{\partial t} \\ \vdots \\ \frac{\partial x_n}{\partial t} \end{array} \right) = \left\langle \nabla f, \frac{\partial r}{\partial t} \right\rangle
\]
We get again that the gradient is orthogonal to the slope of the curve.

### 2.3 Definition of tangent hyperplane

With the definitions and theorems seen before we can now define the tangent hyperplane in a point and get the equation. In the previous section we saw the theorem that the gradient is perpendicular to the level curve or the level surface of a function. If you recall the definition of hyperplanes and how to get one, we need a point and an orthogonal vector. Given any point, now we have a perpendicular vector to the function, which is the gradient.

Consider a function of two variables:

\[
  f : \mathbb{R}^2 \to \mathbb{R} \\
  (x, y) \to z = f(x, y)
\]

We can transform this function into a "level surface", like it was the level surface of a higher dimensional function:

\[
  z = f(x, y) \to F(x, y, z) = f(x, y) - z = 0
\]

Now we treat it as a level surface, so the gradient of the function \( F \) will be orthogonal to any point of the function. Let \( F(x, y, z) \) define a surface. Then the tangent plane to \( F(x, y, z) \) at a point \((x_0, y_0, z_0)\) is the plane with normal vector \( \nabla F(x_0, y_0, z_0) \) that passes through the point \((x_0, y_0, z_0)\). \( F \) must be differentiable at a point \((x_0, y_0, z_0)\). In general, the equation of the tangent plane is:

\[
  \langle \nabla f, r - r_0 \rangle = \frac{\partial F}{\partial x}(x_0, y_0, z_0)(x-x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y-y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z-z_0) = 0
\]

In particular, considering the previous function:

\[
  F(x, y, x) = f(x, y) - z \to \nabla F = \begin{pmatrix}
  \frac{\partial F}{\partial x} \\
  \frac{\partial F}{\partial y} \\
  \frac{\partial F}{\partial z}
\end{pmatrix} = \begin{pmatrix}
  \frac{\partial f}{\partial x} \\
  \frac{\partial f}{\partial y} \\
  -1
\end{pmatrix}
\]

Then we have:

\[
  z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y-y_0) \text{ with } z_0 = f(x_0, y_0)
\]
This is the equation of the tangent plane to surface \( z = f(x, y) \) at point \((x_0, y_0)\).

**Example:** consider the surface \( z = f(x, y) = 3x^2 - xy \). We want to obtain the tangent plane at the point \((1, 2, 1)\).

First, we define the surface as a level surface:
\[ z = 3x^2 - xy \rightarrow F(x, y, z) = 3x^2 - xy - z \]

We calculate the gradient at the given point:
\[ \nabla F = (6x - y - x - 1) \rightarrow \nabla F(1, 2, 1) = (4 \quad -1 \quad -1) \]

Using the scalar product, we write down the equation of the tangent plane:
\[
\langle \nabla F, r - r_0 \rangle = (4 \quad -1 \quad -1) \begin{pmatrix} x - 1 \\ y - 2 \\ z - 1 \end{pmatrix} = 0 \rightarrow 4x - y - z = 1
\]

### 2.4 Example of tangent hyperplanes and applications

#### 2.4.1 Tangent hyperplanes for contour lines

We can also calculate tangent hyperplanes of contour lines. Remember that the gradient is perpendicular to the contour lines, hence the gradient will be also the orthogonal vector of the tangent hyperplane.

**Example:** consider the function \( f(x, y) = x^2 + y^2 \), which represents a paraboloid and its level curve \( f(x, y) = x^2 + y^2 = k \). Find the tangent plane (line in this case) of this level curve: \( f(x, y) = x^2 + y^2 = 8 \) at the point \( p_0 = (2, 2) \).

Calculate the gradient of the function at \((2, 2)\):
\[
\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \rightarrow \nabla f(2, 2) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}
\]

Write the equation of the tangent plane:
\[
z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)
\]
Now $z = 8$ and $z_0 = f(2, 2) = 2^2 + 2^2 = 8$:

$$8 - 8 = (4 \quad 4) \begin{pmatrix} x - 2 \\ y - 2 \end{pmatrix} = 4x + 4y - 16 \rightarrow x + y = 4$$

This can be graphically seen also. The level curve $f(x, y) = x^2 + y^2 = 8$ represents a circumference of radius $\sqrt{8}$. Drawing it and taking into account that the gradient is perpendicular to the tangent vector we have what is shown in Figure 6.

![Tangent hyperplane (tangent line) of the circumference $x^2 + y^2 = 8$. Black line represents the tangent line with equation $x + y = 4$. Blue arrow represents the gradient vector.](image)

Figure 6: Tangent hyperplane (tangent line) of the circumference $x^2 + y^2 = 8$. Black line represents the tangent line with equation $x + y = 4$. Blue arrow represents the gradient vector.

### 2.4.2 Lagrange multipliers

We can recover the definition of the Lagrange multipliers with the level curve and tangent hyperplanes. The idea is that there are points in a function
f(x, y) at which the level curve $f(x, y) = k$ coincide, and is tangent, with the constraint. As the points are the same, the tangent lines are also the same for $f(x, y)$ and for the constraints $G(x, y)$. Hence, to find an optimal point of $f$ subject to the constraint $G(x, y)$ we impose that the gradients must be proportional to each other:

$$\nabla f = \lambda \nabla G \text{ and } G(x, y) = 0$$
3 Exercises

Ex. 1 — Given a plane in $\mathbb{R}^3$,

$$-x + 2z = 10$$

and the line

$$r = \begin{pmatrix} 5 \\ 2 \\ 10 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}$$

Determine whether they are orthogonal, parallel or neither.

Ex. 2 — Consider the surface $x^2yz - y + z = 0$. Obtain the tangent plane at the point $p_0 = (1, 2, 3)$ and the normal line through that point.

Ex. 3 — Find the greatest and smallest values of $f(x, y) = xy$ on the ellipse

$$\frac{x^2}{8} + \frac{y^2}{2} = 1$$

Verify that

$$\nabla f = \lambda \nabla G$$

and the optimal point
References


