Constrained optimization:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad f(x) \\
\text{s.t.} & \quad h_1(x) = 0 \\
& \quad h_2(x) = 0 \\
& \quad \vdots \\
& \quad h_m(x) = 0 \\
\text{subject to} & \quad g_1(x) \leq 0 \\
& \quad g_2(x) \leq 0 \\
& \quad \vdots \\
& \quad g_l(x) \leq 0 
\end{align*}
\]

Example:

\[
\begin{align*}
\min_{x \in \mathbb{R}^2} & \quad x_1^2 + x_2^2 \\
\text{s.t.} & \quad x_1 - x_2 = 1 \\
& \quad x_1^2 - x_2^2 \leq 4
\end{align*}
\]

\[
\Rightarrow f(x) = x_1^2 + x_2^2 \\
& \quad h_1(x) = x_1 - x_2 - 1 \\
& \quad g_1(x) = x_1^2 - x_2^2 - 4
\]

- Let's focus on problems with equality constraints:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad f(x) \\
\text{s.t.} & \quad h_1(x) = 0 \\
& \quad h_2(x) = 0 \\
& \quad \vdots \\
& \quad h_m(x) = 0
\end{align*}
\]

Example:

\[
\begin{align*}
\min_{x \in \mathbb{R}^2} & \quad x_1^2 - x_2^2 \\
\text{s.t.} & \quad x_1 + x_2 = 2 \\
& \quad x_1^2 - x_2^2 - 1 = 0
\end{align*}
\]

\[
\Rightarrow f(x) = x_1^2 - x_2^2, \quad h(x) = \begin{bmatrix} h_1(x) \\ h_2(x) \end{bmatrix}
\]
- $x_*$ is a local min if it is locally the best point.

- $x_*$ is a local point if there exists an $\varepsilon > 0$ such that $x_*$ is the global (best) solution to

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$h(x) = 0$$

$$\|x - x_*\| \leq \varepsilon$$

Example

- How to find local solutions for constrained problems?

- We need the notions of feasible direction and tangent plane.

- Note: $\{x \mid h(x) = 0\}$ is a surface.
Examples:

\[ h(x) = x_1^2 + x_2^2 - 1 \quad \text{and} \quad x \in \mathbb{R}^2 \Rightarrow \]

\[ h(x) = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ x_3 - 1 \end{bmatrix} \quad \text{and} \quad x \in \mathbb{R}^3 \]

\[ \Rightarrow \]

\[ \Rightarrow \ \min f(x) \ \text{s.t.} \ h(x) = 0 \ \text{is equivalent to optimizing a function over a surface} \]

- Regular point: A point \( y \in \mathbb{R}^n \) is called a regular point for the surface \( \{ x \mid h(x) = 0 \} \) if the vectors \( \nabla h_1(y), \nabla h_2(y), \ldots, \nabla h_m(y) \) are linearly independent, i.e., if \( \alpha_1 \nabla h_1(y) + \cdots + \alpha_m \nabla h_m(y) = 0 \) for some \( \alpha_1, \ldots, \alpha_m \) then \( \alpha_1 = \cdots = \alpha_m = 0 \)

- Examples:

\[ x_1^2 + x_2^2 = 1 \Rightarrow h(x) = (x_1^2 + x_2^2 - 1) \]

\[ \Rightarrow \nabla h(y) = [2x_1, 2x_2] \]
If $\lambda_1 \nabla h_1(y) = 0$ and $\lambda_1 \neq 0$, then

$\nabla h_1(y) = 0 \Rightarrow [y_1, y_2] = 0$ but $h_1(0) = 0$

$\Rightarrow$ All points for the surface $x_1^2 + x_2^2 = 1$ are regular.

$\Rightarrow h(x) = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ x_3 - 1 \end{bmatrix}$

$\Rightarrow \nabla h_1(y) = [2y_1, 2y_2, 0]$

$\nabla h_2(y) = [0, 0, 1]$

If $\lambda_1 \nabla h_1(y) + \lambda_2 \nabla h_2(y) = 0$ and $[\lambda_1, \lambda_2] \neq 0$

$\Rightarrow \lambda_1 x_0 + \lambda_2 x_1 = 0 \Rightarrow \lambda_2 = 0 \Rightarrow \nabla h_1(y) = 0$

$\Rightarrow y_1 = y_2 = 0$ but $h_1(y) = 0$

$\Rightarrow$ All points are regular.

$\Rightarrow x_1 = 0$

1. $h(x) = [x_1]$ $\Rightarrow \nabla h_1(y) = [1]$

   Linearly independent

   $\Rightarrow$ All points are regular

2. $h(x) = [x_1^2]$ $\Rightarrow \nabla h_1(y) = [2y_1] = 0$

   (since $h(y) = 0$)

   $\Rightarrow$ No point is regular.
The vertical line has at least two different representations: for one of them, all points are regular; for another one, none of the points are regular.

The notion of regular points depends not only on the surface, but also on its algebraic representation.

Consider a point $x_*$ such that $h(x_*) = 0$.

What are the directions $\Delta x$ such that if you move along them by starting at $x_*$, you don't fall off the surface $\{ x \mid h(x) = 0 \}$ immediately?

The set of all such directions is the tangent plane at $x_*$.

⇒ all these directions are wrong. The direction should be tangent to the circle.
- How to characterize the tangent plane?

**Theorem:** If $x_*$ is a regular point, then the tangent plane at $x_*$ for the surface $\{ x \mid h(x) = 0 \}$ is characterized as:

\[ \{ \Delta x \mid \nabla h(x_*) \Delta x = 0 \} \]

or

\[ \{ \Delta x \mid \nabla h_1(x_*) \Delta x = 0, \ldots, \nabla h_m(x_*) \Delta x = 0 \} \]

**Examples:**

\[ x_1^2 + x_2^2 = 1 \implies \nabla h(x_*) = [2x_1^* \ 2x_2^*] \]

\[ \nabla h(x_*) \Delta x = 0 \implies \text{Tangent Plane: } \{ \Delta x \mid 2x_1^* \Delta x_1 + 2x_2^* \Delta x_2 = 0 \} \]
\[ x_1^2 + x_2^2 = 1 \]

\[ \Rightarrow \nabla h(x_*) = \begin{bmatrix} 2x_1^* & 2x_2^* & 0 \end{bmatrix} \]

\[ \Rightarrow \text{Tangent plane } = \{ \Delta x \ | \ \nabla h(x_*) \Delta x = 0 \} \]

\[ = \{ \Delta x \ | \ x_1^* \Delta x_1 + x_2^* \Delta x_2 = 0 \text{ and } \Delta x_3 = 0 \} \]

Some intuition behind proof (not 100% accurate):

Assume \( h(x_*) = 0 \) and \( h(x_* + \Delta x) = 0 \). Then

\[ h_i(x_* + \Delta x) = h_i(x_*) + \nabla h_i(x_*) \Delta x + \ldots \ldots \text{ } i = 1, \ldots, m \]

\[ \Rightarrow \nabla h_i(x_*) \Delta x = 0 \text{ for small } \Delta x \Rightarrow \nabla h(x_*) \Delta x = 0 \]

\[ \min_{x \in \mathbb{R}^n} f(x) \] subject to \( h(x) = 0 \) \( \Rightarrow \) Assume \( x_* \text{ is local minimum} \)

\[ \Rightarrow \] \( f(x_* + \Delta x) \geq f(x_*) \) for small feasible \( \Delta x \)

\[ \Rightarrow \] \( f(x_* + \Delta x) \geq f(x_*) \) for all small \( \Delta x \) such that \( \Delta x \in \text{Tangent plane at } x_* \).
Assume also \( x^* \) is a regular point.

\[ f(x^* + \Delta x) \geq f(x^*) \text{ for all small } \Delta x \text{ such that } \nabla h(x^*) \Delta x = 0 \]

Also: \[ f(x^* + \Delta x) - f(x^*) = \nabla f(x^*) \Delta x + \ldots \]

Theorem: If \( x^* \) is regular and a local min, then

\[ \nabla f(x^*) \Delta x \geq 0 \text{ for all small } \Delta x \text{ s.t. } \nabla h(x^*) \Delta x = 0 \]

First-order necessary condition: If \( x^* \) is a local min and regular, then there exist constants \( \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R} \) such that

\[ \nabla f(x^*) + \lambda_1 \nabla h_1(x^*) + \ldots + \lambda_m \nabla h_m(x^*) = 0 \]