7. Constrained optimization

1. Feasible direction

(P) minimize $f(x)$

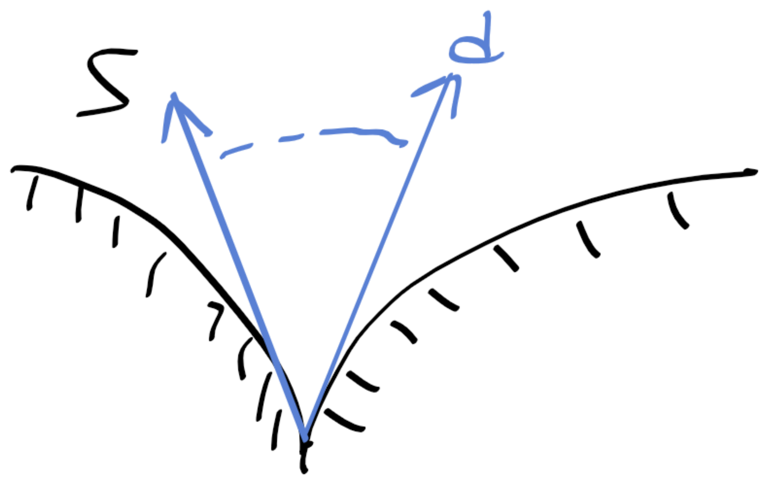
subject to $x \in S = \{x \in X \subseteq \mathbb{R}^n : g_i(x) \le 0, \; i = 1, \dots, m\}$

where we assume that f and $g_i \in C^1(X)$ if X is open, otherwise C^1 in a neighbourhood of X.

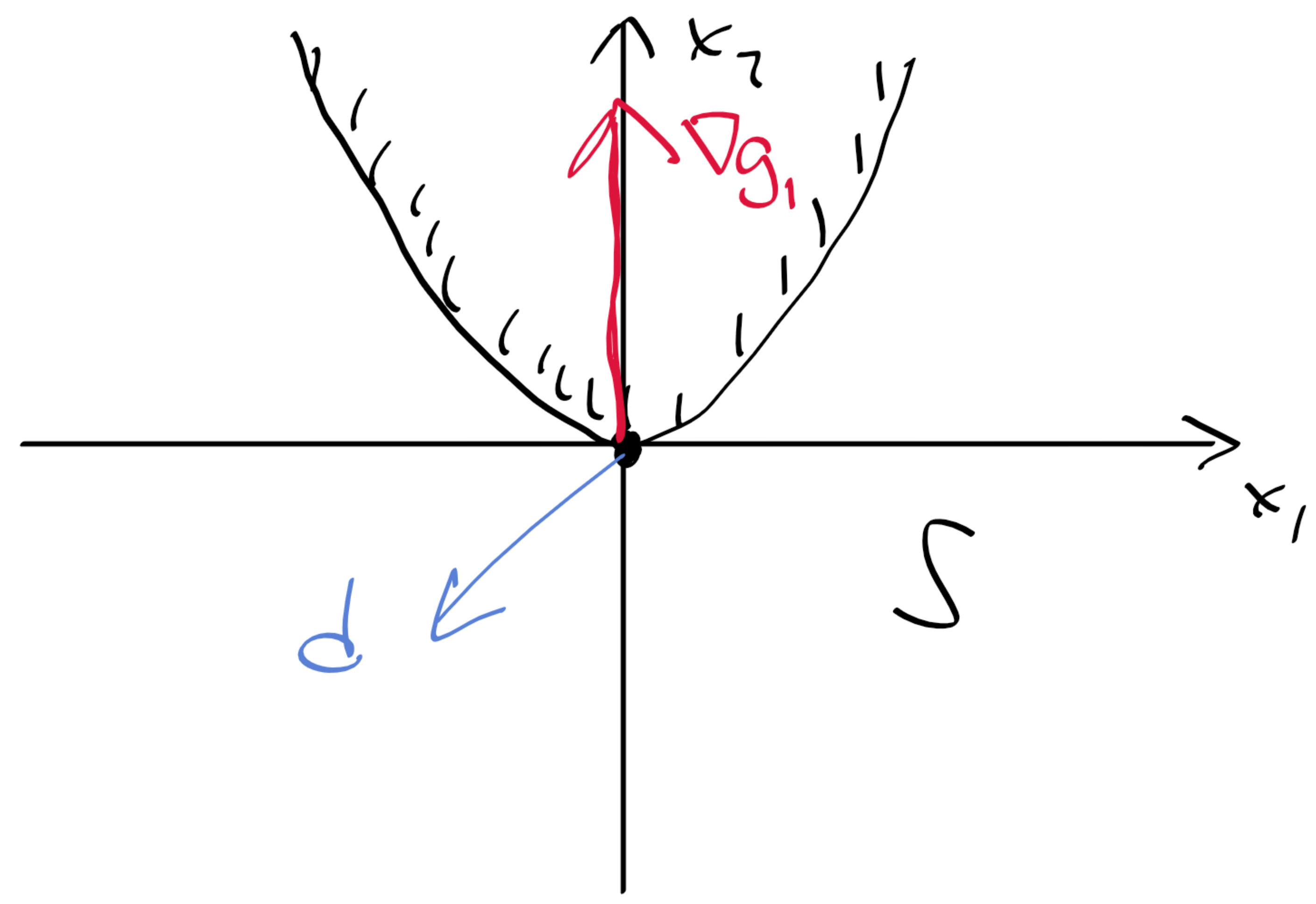
Goal: find necessary conditions for local min.

Def. A direction $d \neq 0$ is feasible at $a \in S$ iff $\exists\, \delta > 0$: $a + td \in S \quad \forall\, 0 < t < \delta$.

Ex.



Ex. $g_1(x) = x_2 - x_1^2 \le 0$

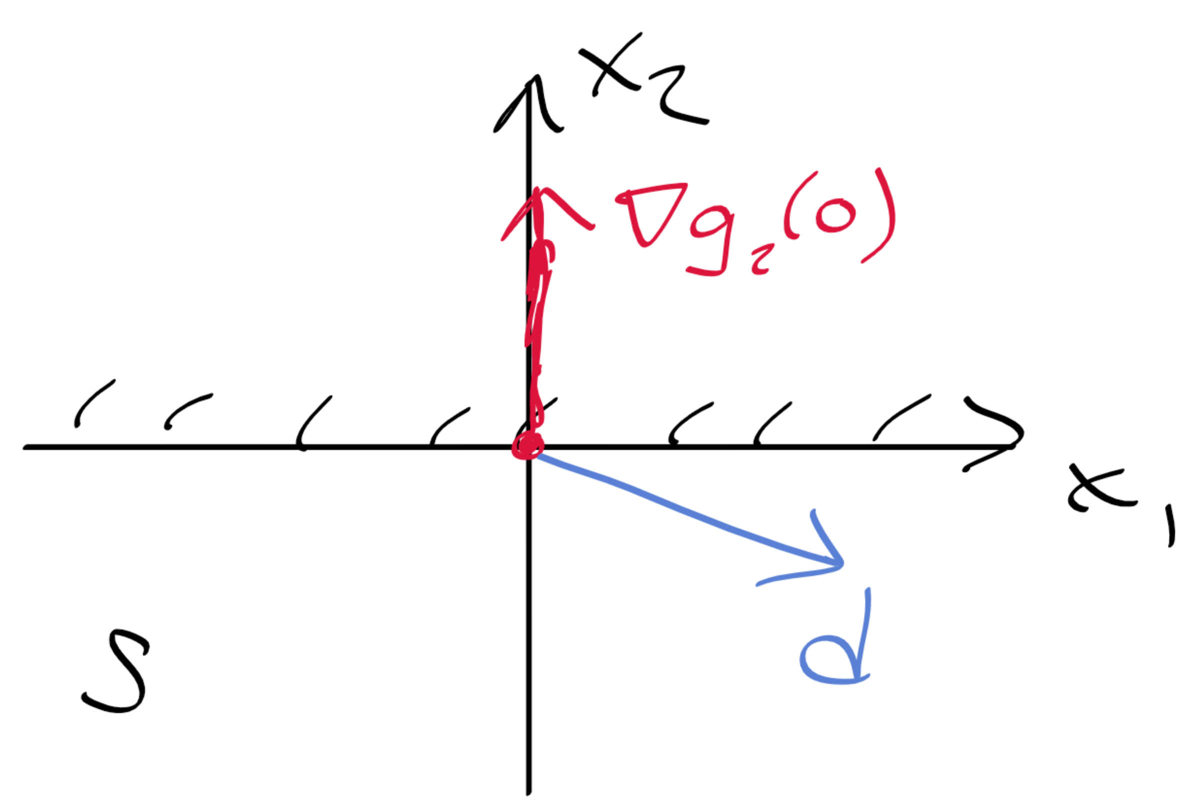


Feasible directions at the origin are $d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ with $d_1 \in \mathbb{R}$, $d_2 \le 0$.

Note $\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix}$, $\nabla g_1(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and the directional derivative is $\nabla g_1(0)^T d = d_2 \le 0$

Ex. $g_2(x) = x_2 \le 0$

The same conclusion!

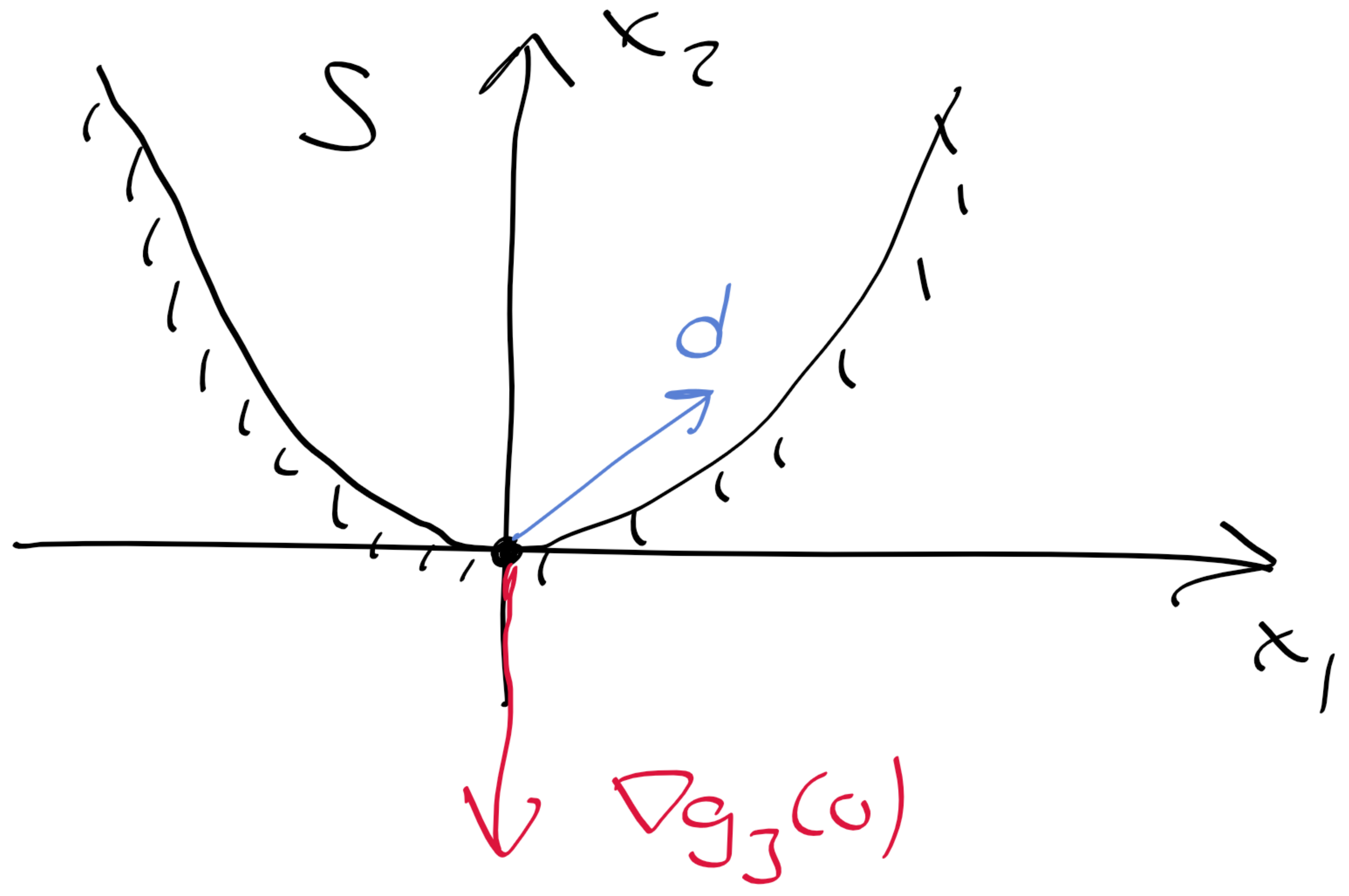


Ex. $g_3(x) = x_1^2 - x_2 \leq 0$

Feasible directions at the origin are

$d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ with $d_1 \in \mathbb{R}$

$d_2 > 0$.



$\nabla g_3(x) = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}$, $\nabla g_3(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

and $\nabla g_3(0)^T d = -d_2 < 0$

Lemma: $\bar{x}$ local minimizer of (P) $\Rightarrow$

$$\nabla f(\bar{x})^T d \geq 0 \quad \forall d \text{ feasible.}$$

Proof: $f(\bar{x} + td) \geq f(\bar{x})$ $\forall d$ feasible, small $t > 0$

$$0 \leq \frac{f(\bar{x}+td) - f(\bar{x})}{t} \to f'(\bar{x}; d) = \nabla f(\bar{x})^T d \quad \text{as } t \to 0^+$$

#

General characterization of a feasible direction d at $\bar{x}$ of a scalar active constraint $g(x) \leq 0$ $(g(\bar{x}) = 0)$:

For small $t > 0$ we want

$$0 \underset{\substack{\uparrow \\ d \text{ is feasible}}}{\geq} g(\bar{x} + td) \underset{\substack{\uparrow \\ \text{Taylor exp.}}}{=} g(\bar{x}) + \nabla g(\bar{x})^T (td) + o(t)$$

$$= 0 + t\big(\nabla g(\bar{x})^T d + \underbrace{o(1)}_{\to 0 \text{ as } t \to 0^+}\big)$$

Lemma: $\nabla g(\bar{x})^T d < 0 \Rightarrow$

d is a feasible direction at $\bar{x}$ and $\nabla g(\bar{x}) \neq 0$

7 Constrained optimization

2. KKT-theory for inequality constraints

(P) minimize $f(x)$
subject to $x \in S = \{x \in \mathbb{X} \subseteq \mathbb{R}^n : g_i(x) \leq 0, \ i = 1, \dots, m\}$

A necessary condition for $\bar{x}$ to be a local minimizer, where the active constraints are $\{g_i(x) \leq 0, i = 1, \dots, p \leq m\}$, is:

$$(1) \quad \begin{cases} \nabla g_i(\bar{x})^T d < 0 \\ i = 1, \dots, p \end{cases} \implies \nabla f(\bar{x})^T d \geq 0$$

From this we want to derive conditions without d.

Let

$$A = \begin{pmatrix} \nabla f(\bar{x})^T \\ \nabla g_1(\bar{x})^T \\ \vdots \\ \nabla g_p(\bar{x})^T \end{pmatrix}. \quad \text{Then } (1) \iff$$

(2) $\quad Ad < 0$ has no solution

We rewrite this by means of Farkas' Theorem:

Farkas' Theorem

Exactly one of $\boldsymbol{Ax} \leq \boldsymbol{0}$, $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} > 0$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{c}$, $\boldsymbol{y} \geq \boldsymbol{0}$ has a solution.

In (2) we will "close the gap to zero" in the component $(Ad)_k$ of the column vector Ad that is closest to zero: Set

$$\alpha := \min_k |(Ad)_k|. \quad \text{Then } (2) \iff$$

$$Ad + \alpha \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \leq 0, \quad \alpha > 0 \quad \text{has no solution} \iff$$

$$\begin{pmatrix} A & \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{pmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix} \leq 0, \quad \begin{pmatrix} 0_{1 \times (p+1)} & 1 \end{pmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix} > 0 \quad \text{has no sol.}$$

Farkas' thm
$\iff$

$$\begin{pmatrix} A^T \\ (1 \ \cdots \ 1) \end{pmatrix} y = \begin{pmatrix} 0_{(p+1)\times 1} \\ 1 \end{pmatrix} \quad \text{has a solution}$$

$$y \geq 0$$

$$\iff \begin{cases} A^T y = 0 \\ \sum y_i = 1 \\ y \geq 0 \end{cases} \text{has sol.} \iff \begin{cases} A^T(\beta y) = 0 \\ \sum \beta y_i = \beta > 0 \text{ has sol.} \\ \beta y \geq 0 \end{cases}$$

$$u := \beta y \quad \iff \begin{cases} A^T u = 0 \\ u \neq 0 \\ u \geq 0 \end{cases} \text{has sol.} \iff$$

$$(3) \quad \exists u: \begin{cases} \begin{pmatrix} \nabla f(\bar{x}) & \nabla g_1(\bar{x}) \ \cdots \ \nabla g_p(\bar{x}) \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_p \end{pmatrix} = 0 \\ u \geq 0 \\ u \neq 0 \end{cases}$$

To include the inactive constraints ($g_i(\bar{x}) < 0$, $i = p+1, \dots, m$)

we define $u_i = 0$ for $i = p+1, \dots, m$.

Then $u_i g_i(\bar{x}) = 0 \quad \forall i = 1, \dots, m$.

(3) $\iff$ one of the following is true:

- $u_0 \neq 0$: (we can assume that $u_0 = 1$)

$$(KKT) \quad \begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0 \\ u_i g_i(\bar{x}) = 0 \quad \forall i \\ u_i \geq 0 \quad \forall i \\ g_i(\bar{x}) \leq 0 \end{cases}$$

- $\underline{u_0 = 0}$: Then
$$\begin{cases} \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0 \\ u_i g_i(\bar{x}) = 0 \quad \forall i \\ u_i \geq 0 \quad \forall i \\ \text{Some } u_i \neq 0 \end{cases}$$

i.e. the gradients of the active constraint $\{\nabla g_i(\bar{x}), i=1,\dots,p\}$ are positively linearly dependant (PLD). The converse is positively linerly independant (PLI).

Def. (Constraint qualification)

The gradients of the active constraints $\{\nabla g_i(\bar{x}),\ i=1,\dots,p\}$ are PLI $\iff$

$$(CQ) \quad \begin{cases} \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0 \\ u_i g_i(\bar{x}) = 0 \\ u_i \geq 0 \end{cases} \Rightarrow \text{all } u_i = 0$$

Thm 3: $\bar{x}$ satisfies (CQ) and solves (P)

$$\Rightarrow \quad \exists u: \begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0 \\ u_i g_i(\bar{x}) = 0 \\ u_i \geq 0 \end{cases}$$