7.4 Sufficient conditions for a minimum

1. First-order conditions

$$(P) \quad \begin{array}{l} \text{minimize } f(x) \\ \text{subject to } x \in S = \{x \in X : g(x) \leq 0, h(x) = 0\} \end{array}$$

$$\mathcal{L}(x; u, v) = f(x) + u^T g(x) + v^T h(x)$$

$$I(\bar{x}) = \{ i : g_i(\bar{x}) = 0 \}$$

Thm 5. If $f, g_i \in C^1(X)$, $i \in I(\bar{x})$ are convex and h_j affine in a neighbourhood of a KKT point $(\bar{x}; \bar{u}, \bar{v})$ with $\bar{x} \in S$, then $\bar{x}$ is a local minimizer.

Proof: $\mathcal{L}(\cdot; \bar{u}, \bar{v}) = f + \underbrace{u^T g}_{\text{sum is convex}} + \underbrace{v^T h}_{\text{sum is convex}}$

is convex. (KKT) $\Rightarrow$ $\nabla_x \mathcal{L}(\bar{x}; \bar{u}, \bar{v}) = 0$
so $\mathcal{L}(\cdot; \bar{u}, \bar{v})$ has a local min. at $\bar{x}$:

$$\mathcal{L}(\bar{x}; \bar{u}, \bar{v}) \leq \mathcal{L}(x; \bar{u}, \bar{v}) \quad \forall x \in S \text{ close to } \bar{x}$$

LHS: $\mathcal{L}(\bar{x}; \bar{u}, \bar{v}) = f(\bar{x}) + \underbrace{\bar{u}^T g(\bar{x})}_{=0 \text{ (KKT)}} + \bar{v}^T \underbrace{h(\bar{x})}_{=0 \text{ in } S} = f(\bar{x})$

RHS: $\mathcal{L}(x; \bar{u}, \bar{v}) = f(x) + \underbrace{\underbrace{\bar{u}^T}_{\geq 0} \underbrace{g(x)}_{\leq 0}}_{\leq 0} + \bar{v}^T \underbrace{h(x)}_{\substack{=0 \\ \text{in } S}} \leq f(x)$

Hence

$$f(\bar{x}) \leq f(x) \quad \forall x \in S \text{ close to } \bar{x} \quad \#$$

Cor. If we have also global convexity of f and all g_i on $\underline{X}$ convex, then $\bar{x}$ is a global minimizer.

Remark: We need not check (CQ) !

2a. Sufficient conditions for minimum

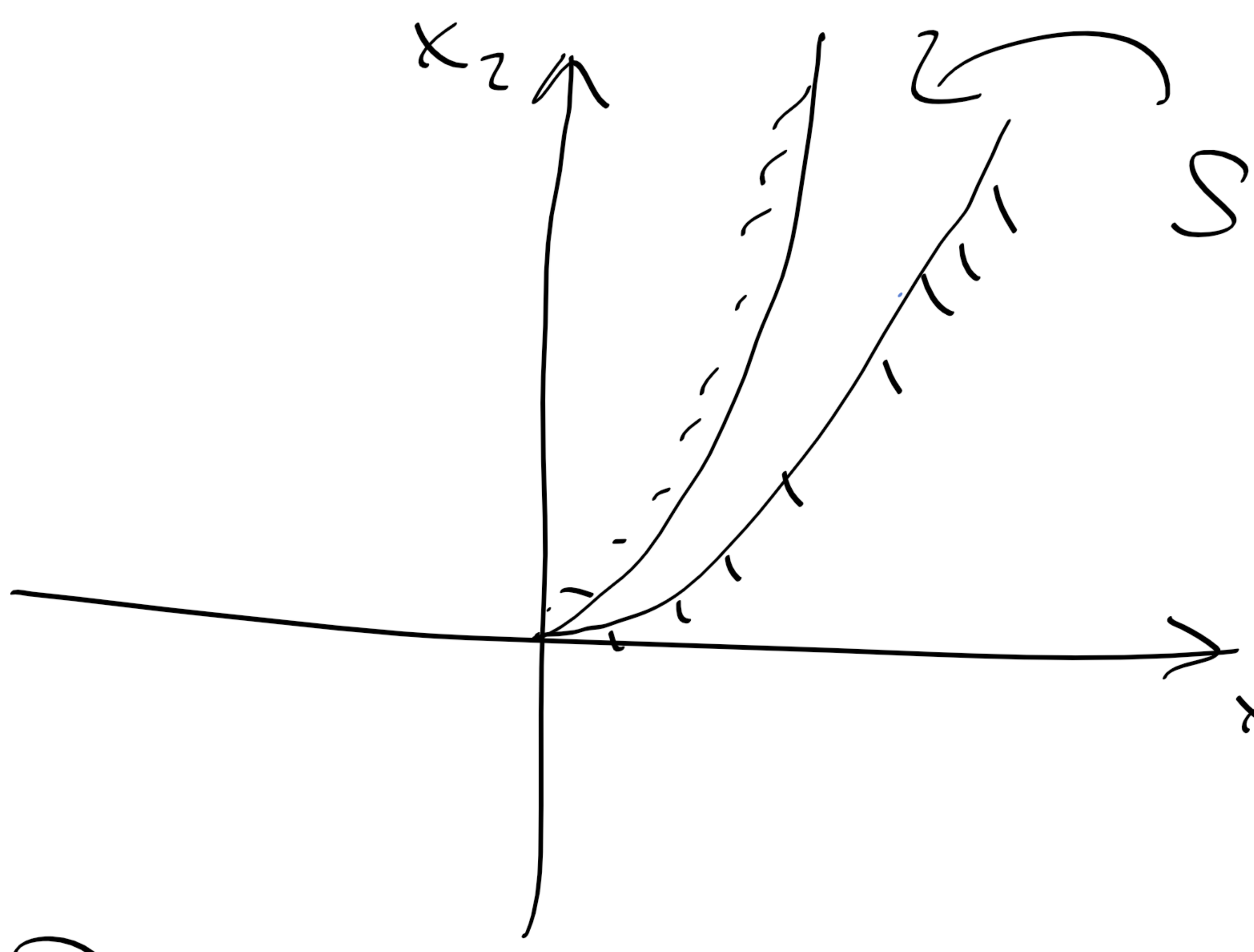
The idea is to generalize the unconstrained

$$\underset{x \in \mathbb{R}^n}{\text{minimize}}\ f(x) \quad :$$

$\nabla f(\bar{x}) = 0$, $\nabla^2 f(\bar{x})$ pos. def. $\Rightarrow$ $\bar{x}$ strict local min.

and use $\mathcal{L}(\cdot\,; u, v)$.

The previously defined set of feasible directions is too small, for example,



$S = \{(x_1, x_2) : x_1^2 \le x_2 \le 2x_1^2,\ x_1 \ge 0\}$

has no feasible direction at the origin.

Def. Feasible directions for the linerized active constraints $\tilde{C}(\bar{x}) = \{ d \neq 0 : \nabla g_i^T d \le 0,\ i \in I(\bar{x}),\ \nabla h_j^T d = 0\ \forall j \}$

In the figure at $\bar{x} = 0$: $d \in \tilde{C}(\bar{x}) \Leftrightarrow$

$d = \begin{pmatrix} d_1 \\ 0 \end{pmatrix}$, $\cancel{d_1 > 0}$ (check this!)

Thm 6. (Second-order sufficient conditions)

Assume $f, g_i, h_j \in C^2(\bar{X})$ and that $\bar{x} \in S$, $\bar{u}, \bar{v}$ satisfy

- (KKT)
- $\nabla_x^2 \mathcal{L}(\bar{x}; \bar{u}, \bar{v})$ is pos. def. $\forall d \in \tilde{C}(\bar{x})$

Then $\bar{x}$ is a strict local minimizer.

Proof: Let $\{x_k\}_1^\infty$ be a sequence in S with $x_k \neq \bar{x}$ $\forall k$ and $x_k \to \bar{x}$. Set

$$\lambda_k := \| x_k - \bar{x} \| \quad \text{and} \quad d_k := \frac{x_k - \bar{x}}{\lambda_k}$$

Since $\|d_k\| = 1$ (compact set), Bolzano-Weierstraß then gives a convergent subsequence $\{d_k\}_{k \in \mathcal{K} \subseteq \mathbb{N}}$ with $d_k \to d$ as $k \to \infty$.

Thus $x_k = \bar{x} + \lambda_k d_k$ with $\lambda_k \to 0$, $k \to \infty$.

$d \in \tilde{C}(\bar{x})$ because $(i \in I(\bar{x}))$

- $$0 \geq g_i(x_k) \underset{\text{Taylor}}{=} g_i(\bar{x}) + \nabla g_i(\bar{x})^T (\lambda_k d_k) + \mathcal{O}(\lambda_k^2)$$
  $$= 0 + \underbrace{\lambda_k}_{>0} \left(\nabla g_i(\bar{x})^T d_k + \mathcal{O}(\lambda_k) \right)$$
  $k \to \infty \Rightarrow \nabla g_i(\bar{x})^T d \leq 0$
- Similarly for $h_j(\bar{x})$

Then $\quad f(x_k) \geq f(x_k) + \underbrace{\bar{u}^T g(x_k)}_{\leq 0} + \bar{v}^T \underbrace{h(x_k)}_{=0}$

$= \mathcal{L}(x_k; \bar{u}, \bar{v}) = \mathcal{L}(\bar{x} + \lambda_k d_k; \bar{u}, \bar{v}) = [\text{Taylor exp.}]$

$= \mathcal{L}(\bar{x}; \bar{u}, \bar{v}) + \lambda_k \nabla_x \mathcal{L}(\bar{x}; \bar{u}, \bar{v})^T d_k$

$\quad + \lambda_k^2 d_k^T \nabla_x^2 \mathcal{L}(\bar{x}; \bar{u}, \bar{v}) d_k + o(\lambda_k^2)$

$= f(\bar{x}) + \lambda_k \cdot 0 + \lambda_k^2 \Big(\underbrace{d_k^T \nabla^2 \mathcal{L}(\bar{x}; \bar{u}, \bar{v}) d_k}_{>0} + o(\lambda_k) \Big)$

$> f(\bar{x})$ for large k #

Cor: If instead $u \leq 0$ and

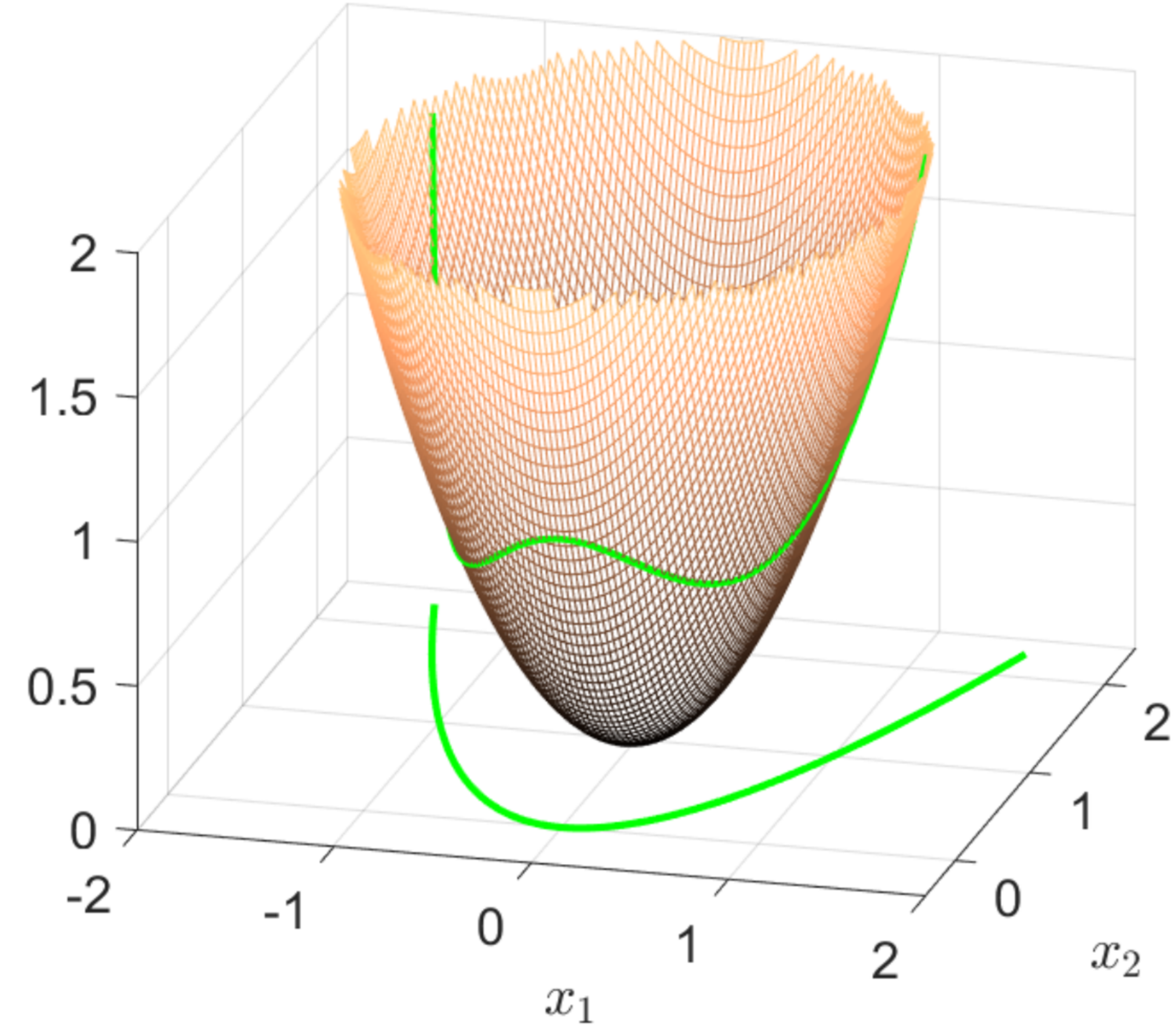
$\nabla_x^2 \mathcal{L}(\bar{x}; \bar{u}, \bar{v})$ neg. def.

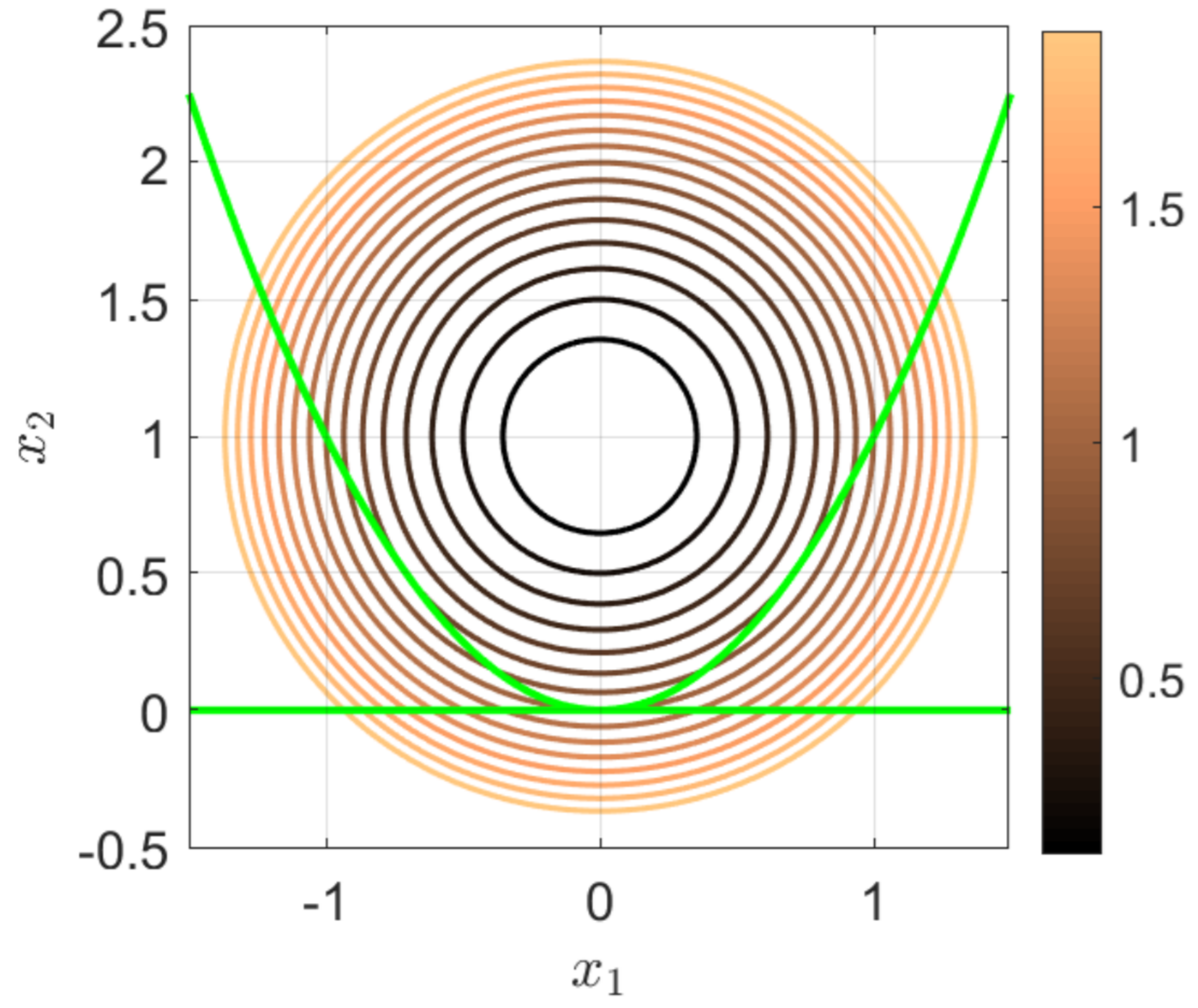
in Thm 6, then $\bar{x}$ is a strict local maximizer.

2b. Second-order sufficient conditions: improvement

Previous example:

minimize $f(x_1, x_2) = x_1^2 + (x_2 - 1)^2$ subject to $0 \le x_2 \le x_1^2$





$$g_1(x) = x_2 - x_1^2 \le 0$$
$$g_2(x) = -x_2 \le 0$$

$$\nabla f = \begin{pmatrix} 2x_1 \\ 2(x_2 - 1) \end{pmatrix}, \quad \nabla g_1 = \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix}, \quad \nabla g_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\mathcal{L} = f + u_1 g_1 + u_2 g_2$$

$$\nabla_x \mathcal{L} = \nabla f + u_1 \nabla g_1 + u_2 \nabla g_2 = \begin{pmatrix} 2x_1 - u_1 2x_1 \\ 2(x_2 - 1) + u_1 - u_2 \end{pmatrix}$$

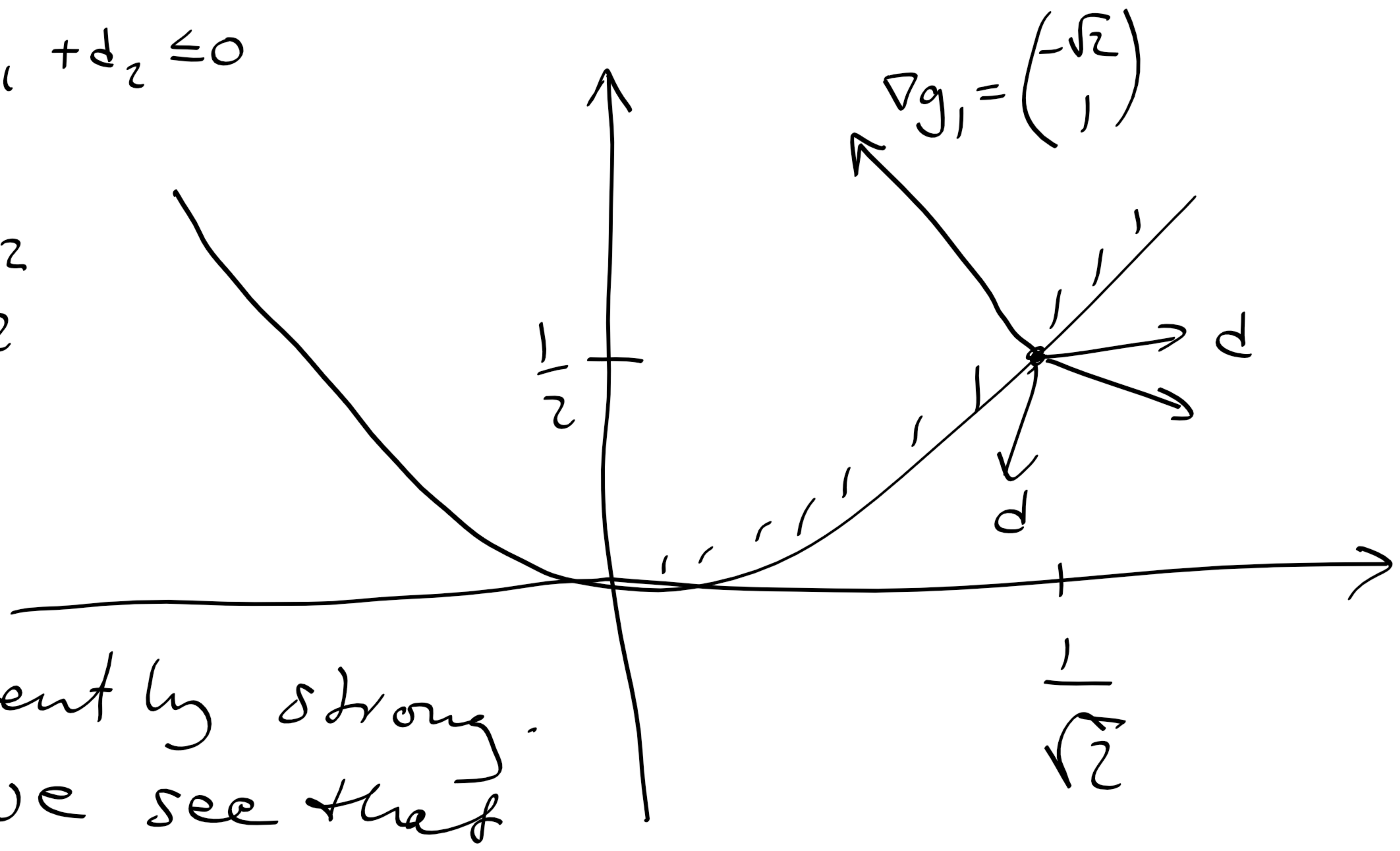
$$\nabla_x^2 \mathcal{L} = \begin{pmatrix} 2(1 - u_1) & 0 \\ 0 & 2 \end{pmatrix}$$

$$(KKT) \quad \begin{cases} x_1 = (\pm)\frac{1}{\sqrt{2}}, & x_2 = \frac{1}{2} \\ u_1 = 1, & u_2 = 0 \end{cases}$$

only g_1 is active

Directions $d \in \tilde{C}(\bar{x}) = \{ d \ne 0 : \nabla g_1(\bar{x})^T d \le 0 \}$ satisfy

$$\begin{pmatrix} -\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \le 0 \iff -\sqrt{2} d_1 + d_2 \le 0$$



$$d^T \nabla_x^2 \mathcal{L}\left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{2} \end{pmatrix}, 1, 0\right) d = 2d_2^2$$

only pos. semidef.

e.g. $d_1 = 1$, $d_2 = 0$

Thm 6 is not sufficiently strong.
For many d at $\bar{x}$, we see that f increases.

Thm. Given a KKT point $(\bar{x}; \bar{u}, \bar{v})$ and $d \in \tilde{C}(\bar{x})$.

$$\bar{u}_k > 0, \quad \nabla g_k(\bar{x})^T d < 0 \implies \nabla f(\bar{x})^T d > 0$$

i.e. in a direction d towards $\text{int}(S)$, the function values of f increase strictly.

Proof: KKT-point:

$$\nabla f(\bar{x}) + \sum u_i \nabla g_i(\bar{x}) + \sum v_j h_j(\bar{x}) = 0$$

$$\implies \quad \nabla f(\bar{x})^T d + \sum \underbrace{u_i}_{\geq 0} \underbrace{\nabla g_i(\bar{x})^T d}_{\leq 0} + \sum v_j \underbrace{h_j(\bar{x})^T d}_{= 0} = 0$$

If $\bar{u}_k > 0$ and $\nabla g_k(\bar{x})^T d < 0 \implies \nabla f(\bar{x})^T d > 0$ #

Consequently, we need only check the second-order conditions in the tangent directions when $\bar{u}_k > 0$.

Divide the index set of active constraints

$$I(\bar{x}) = I^+(\bar{x}) \cup I^0(\bar{x}) \quad \text{where}$$

$$I^+(\bar{x}) := \{ i \in I(\bar{x}) : \bar{u}_i > 0 \}$$

$$I^0(\bar{x}) := \{ i \in I(\bar{x}) : \bar{u}_i = 0 \}$$

Thm 7 (improved version of thm 6)

- KKT point $(\bar{x}; \bar{u}, \bar{v})$
- $d^T \nabla_x^2 L(\bar{x}; \bar{u}, \bar{v}) \, d > 0 \quad \forall d \neq 0$ with

$$\nabla g_i(\bar{x})^T d = 0, \quad i \in I^+(\bar{x}),$$

$$\nabla g_i(\bar{x})^T d \leq 0, \quad i \in I^0(\bar{x})$$

$$\nabla h_j(\bar{x})^T d = 0, \quad \forall j$$

$\implies \bar{x}$ strict local minimizer.

Ex. (again): KKT point $\begin{pmatrix}\bar{x}_1\\\bar{x}_2\end{pmatrix} = \begin{pmatrix}1/\sqrt{2}\\1/2\end{pmatrix}$, $u_1 = 1$, $u_2 = 0$

directions given by $\nabla g_1(\bar{x})^T d = 0 \iff$

$-\sqrt{2}\, d_1 + d_2 = 0$ and $d \neq 0 \Rightarrow$ both $d_1 \neq 0$ and $d_2 \neq 0$

$\Rightarrow \quad d^T \nabla_x^2 \mathcal{L}(\bar{x}; 1, 0)\, d = 2 d_2^2 > 0$

Thm 7 gives $\bar{x}$ strict local minimizer.

7.5 Quadratic Programming

minimize $\quad f(x) = \frac{1}{2} x^T H x + c^T x$

subject to $\quad Ax = b$

$\quad\quad\quad\quad\quad A'x \leq b'$

where A $m \times n$, $m < n$ and $\operatorname{rank} A = m$

- H pos. semidef. $\Rightarrow$ convex problem $\Rightarrow$ global sol.
- H pos. def. $\Rightarrow$ unique global sol.
- KKT - theory (special methods exist)
- If only equality constraints; solve

  $\underset{\text{underdetermined}}{Ax = b} \iff x = Yb + Z \begin{pmatrix} t_1 \\ \vdots \\ t_{n-m} \end{pmatrix}$ ↖ parameters

  and $\underset{t \in \mathbb{R}^{n-m}}{\text{minimize}} \; f(Yb + Zt)$

  without constraints.
- Use duality (Ch. 8)