

d_1 and d_2 are conjugate if

$$d_1^T Q d_2 = 0$$

$Q = \bar{L} \Rightarrow d_1, d_2$ are orthogonal

Suppose : $Q = Q' > 0$: d_1, d_2, \dots, d_k $\Rightarrow d_i^T Q d_j = 0 \quad i \neq j$
 $\Rightarrow d_i^T Q d_i > 0$

- The directions of d_1, \dots, d_k are linearly independent.

Suppose d_1, \dots, d_k such that

$$\alpha_1 d_1 + \dots + \alpha_k d_k = 0$$

$$(x Q) \alpha_1 Q d_1 + \alpha_2 Q d_2 + \dots + \alpha_k Q d_k = 0$$

$$(x d_i) \alpha_1 \underbrace{d_i^T Q d_1}_{=0} + \alpha_2 \underbrace{d_i^T Q d_2}_{=0} + \dots + \alpha_k \underbrace{d_i^T Q d_k}_{=0} = 0$$

- There are at most n mutually independent directions

Let

$$f = \frac{1}{2} x' Q x + c' x + d \quad Q = Q' > 0$$

f has a global min $x^* = -Q^{-1}c$

The conjugate direction method

$$x_0, \underbrace{d_0 \dots d_{n-1}}_{\text{conjugate}}$$

$$x_{k+1} = x_k + \alpha_k d_k$$

α_k performs a exact line search along d_k

$$\rightarrow \alpha_k = - \frac{\nabla' f(x_k) d_k}{d_k' Q d_k}$$

The sequence $\{x_k\} = \{x_0 \dots x_n\}$ is such that

$$x_n = x^* = -Q^{-1}c$$

Define a new algorithm such that
 d_0, d_1, \dots, d_{n-1} are Q-conjugates.
 but not Q-priori complicated.

$$x_0, d_0 := -\nabla f(x_0)$$

$$x_{k+1} = x_k + \alpha_k d_k$$

$$\alpha_k = - \frac{\nabla f(x_k)' d_k}{d_k' Q d_k} \quad \nearrow \alpha^*$$

$$d_{k+1} := -\nabla f(x_{k+1}) + \beta_k d_k \quad \text{pick } d_k, d_{k+1} \text{ Q-conjugate}$$

$$\text{select } \beta_k \text{ such that } \overbrace{d_k' Q d_{k+1}}^{\text{Q-conjugate}} = 0$$

$$d_k' Q d_{k+1} = -d_k' Q \nabla f(x_{k+1}) + \beta_k d_k' Q d_k$$

$$\beta_k = \frac{d_k' Q \nabla f(x_{k+1})}{d_k' Q d_k}$$

$$x_0, d_0 = -\nabla f(x_0)$$

$$x_{k+1} = x_k + \alpha_k d_k$$

$$d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k$$

exact line search

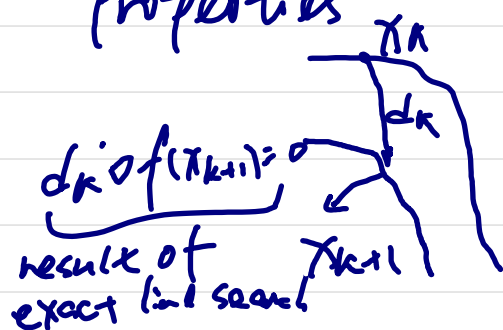
$$\alpha_k = \frac{\nabla f(x_k)' d_k}{d_k' Q d_k}$$

$$\beta_k = \frac{\nabla f(x_{k+1})' Q d_k}{d_k' Q d_k}$$

Q-conjugacy

This algorithm is such that, for any x_0 , the sequence $\{x_k\}$ converges to $x_0 - Q^{-1}c$ in at most n steps, and the directions d_0, \dots, d_{n-1} are mutually Q-conjugate.

Properties



$$\nabla f(x_{k+1})' d_k = 0$$

$$\nabla f(x_k)' \nabla f(x_k) = -\nabla f(x_k)' d_k$$

$$d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k \quad \times (\nabla f'(x_{k+1}))$$

$$\nabla f(x_{k+1})' d_{k+1} = -\nabla f(x_{k+1})'$$

$$f = \frac{1}{2} x' Q x + c' x + d \quad Q = Q' \succ 0$$

$$\nabla f = Qx + c$$

$$x_{k+1} = x_k + \alpha_k d_k$$

$$Qx_{k+1} = Qx_k + \alpha_k Qd_k$$

$$\underbrace{Qx_{k+1} + c}_{\nabla f(x_{k+1})} = \underbrace{Qx_k + c}_{\nabla f(x_k)} + \alpha_k Qd_k$$

$$Qd_k = \frac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k}$$

$$\rho_k = \frac{\nabla f(x_{k+1}) \left[\frac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k} \right]}{d_k' \left[\frac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k} \right]}$$

$$\chi_0.d_0 = -\nabla f(\chi_0)$$

$$\chi_{k+1} = \chi_k + \alpha_k d_k$$

$$d_{k+1} = -\nabla f(\chi_{k+1}) + \beta_k d_k$$

$$\beta_k = \nabla f(\chi_{k+1}) \frac{(\nabla f(\chi_{k+1}) - \nabla f(\chi_k))}{d_k [\nabla f(\chi_{k+1}) - \nabla f(\chi_k)]}$$

χ_k is a line search parameter obtained using a sufficiently accurate line search algorithm.

$$= \frac{\nabla f(\chi_{k+1}) [\nabla f(\chi_{k+1}) - \nabla f(\chi_k)]}{-d_k' \nabla f(\chi_k)}$$

Quasi-Newton

$$x_{k+1} = x_k - \underbrace{[\nabla^2 f(x_k)]^{-1}}_{\substack{\downarrow n}} \nabla f(x_k) \quad \text{Newton}$$

$$x_{k+1} = x_k - \alpha_k \underbrace{[\nabla^2 f(x_k)]^{-1}}_{??} \nabla f(x_k) \quad \rightarrow \text{line search}$$

$$x_{k+1} = x_k - \alpha_k H_k \nabla f(x_k)$$

$$H_{k+1} = ??$$

$$\{H_k\} \rightarrow [\nabla^2 f(x_k)]^{-1}$$

an approximation.

in some sense. $[\nabla^2 f(x_k)]^{-1}_{x=x_k}$

$H_k > 0$ for all k

$$f = \frac{1}{2} x' Q x + c' x + d$$

$$\nabla f = Qx + c \quad \nabla^2 f = Q$$

$$\begin{aligned} \nabla f(y) - \nabla f(x) &= (Qy + c) - (Qx + c) \\ &= Q(y - x) \\ &\quad \downarrow \nabla^2 f \end{aligned}$$

$$Q^{-1}[\nabla f(y) - \nabla f(x)] = y - x$$

$$Q^{-1}[\nabla f(x_{k+1}) - \nabla f(x_k)] = (x_{k+1} - x_k)$$

$$\underline{H_{k+1}} [\underbrace{\nabla f(x_{k+1})}_{\sigma_k} - \underbrace{\nabla f(x_k)}_{\delta_k}] = \underbrace{(x_{k+1} - x_k)}_{\delta_k}$$

$$H_{k+1} \sigma_k = \delta_k$$

n equations in n^2 unknowns

however $H_{k+1} = H_{k+1}'$

The quasi-Newton eqs has several solutions.

Given H_0 , we would like an update law for H
like $H_{k+1} = \dots$

$$H_{k+1} \sigma_k = \delta_k$$

$$H_0 = I$$

$$H_{k+1} = H_k + \frac{\delta_k \delta_k'}{\delta_k' \sigma_k} - \frac{H_k \sigma_k \sigma_k' H_k}{\sigma_k' H_k \sigma_k}$$

outer product inner product

H_0 is symmetric

if H_k is symmetric, then H_{k+1} is also symmetric.

If $\mu_k > 0$, it is always possible to select α_k such that $H_{k+1} > 0$

$$H_{k+1} \sigma_k = H_k \sigma_k + \frac{\delta_k f' \sigma_k}{\delta_k \sigma_k} - \frac{H_k \sigma_k \sigma_k' H_k}{\sigma_k' H_k \sigma_k}$$

$$= H_k \sigma_k + f_k - H_k \sigma_k = f_k$$

Quasi-Newton (x_0, f_0 and f' given) $H_0 = H'_0 > 0$

$$x_{k+1} = x_k - \alpha_k (H_k \sigma f(x_k))$$

$$H_{k+1} = H_k + \frac{\delta_k^* \delta_k'}{\delta_k' \sigma_k} - \frac{H_k \sigma_k \sigma_k' H_k}{\sigma_k' H_k \sigma_k}$$

$$\text{If } f = \frac{1}{2} x' Q x + c' x + d \quad Q = Q', > 0$$

then $\{x_k\} \rightarrow x_* = -Q^{-1}c$ in at most n steps

$\{H_k\} \rightarrow Q^{-1}$ in at most n steps

(provided $\alpha_k = \alpha^*$).

For non-quadratic functions under some assumptions, the line-search is self-accelerating and have a global convergence to $x_0 \in \mathcal{N}$ with quadratic/superlinear speed.

Moreover if $\nabla^2 f(x_k) \succ 0$ then $\{H_k\} \rightarrow [\nabla^2 f(x_k)]^{-1}$

Methods without derivatives

$$x_{k+1} = x_k + \alpha_k d_k$$

α_k is selected using parabolic line search.

$$d_k, d_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, d_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, d_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \dots d_{k-1} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Coordinate directions method

If $\{f(x_k)\}$ is compact then $\{x_k\}$ is such that
 $\lim_{k \rightarrow \infty} x_k \in \mathcal{N}$

$\{x_k\}$ has a limit if in addition $\lim_{k \rightarrow \infty} \|x_{k+1} - x_k\| = 0$

The stopping condition relies on the line search method.

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

guess/pick 3 points

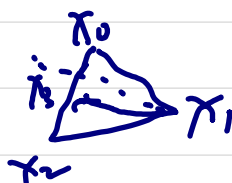
$$\{x_0, x_1, x_2\}$$

↓ projection

$$\{x_0, x_2, x_3\} \quad f(x_3) < f(x_1) \stackrel{?}{=}$$

↓

$$\{x_2, x_3, x_4\}$$



$$f(x_1) > f(x_0) > f(x_2)$$

Simplex method:

n points in n-dimensional space.
discard the worst.
Can cycle between points.

Given P , $\min_x f(x) \quad g(x) \geq 0$ \Leftrightarrow $\min_x f(x) \quad \begin{matrix} g(x) \leq 0 \\ -g(x) \leq 0 \end{matrix}$

n decision variables \Leftrightarrow n decision variables

m equation constraints \Leftrightarrow $2m$ inequality constraints

Given P , $\min_x f(x) \quad h(x) \leq 0$ \Leftrightarrow $\min_x f(x) \quad \begin{cases} h_1(x) + y_1^2 \geq 0 \\ \vdots \\ h_p(x) + y_p^2 \geq 0 \end{cases} \quad h(x) + \gamma^2 = 0$

n decision variables \Leftrightarrow $n+p$ decision variables

p inequality constraints \Leftrightarrow p decision variables

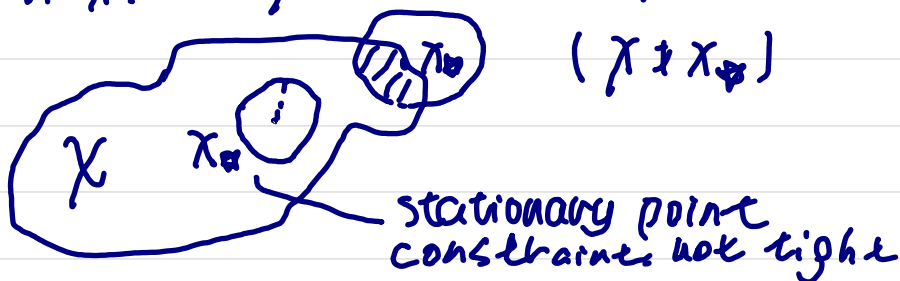
$\min_x f(x)$
 $x \in X$

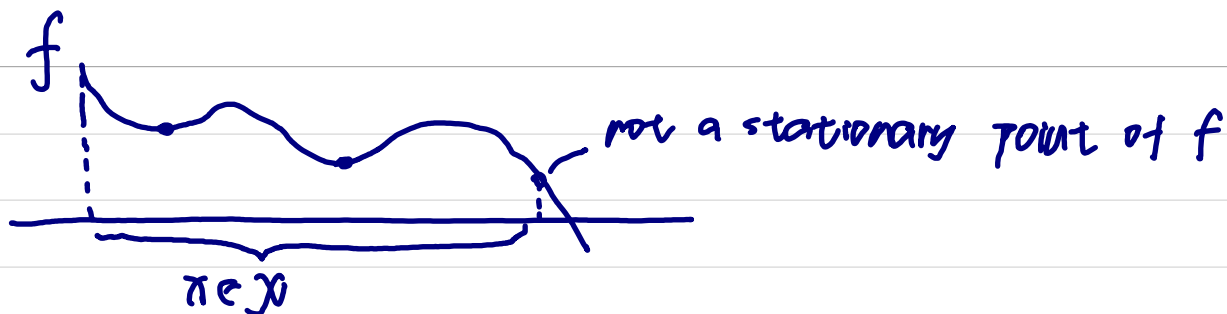
strict

A point x_0 is a local minimiser for the convex programming problem if there exists $\theta (\theta > 0)$ such that

$$f(x_0) \leq f(x) \quad (f(x_0) = f(x))$$

for all $x: x \in X \cap \{ \|x - x_0\| < \theta \}$





problem P_0 (canonical form)

$$\min_x f(x)$$

$$g(x) = 0$$

$$h(x) \leq 0$$

$$h(x) = \begin{cases} h_1(x) \\ \vdots \\ h_p(x) \end{cases}$$

The i -th inequality constraint is active at a point \hat{x} if $h_i(x) = 0$. If $h_i(\hat{x}) < 0$ then the constraint is not active.

All equality constraints are active at any admissible point \hat{x} .

Given P_0 and \hat{x} admissible we define the set $L(\hat{x})$ as the set of all active constraints at \hat{x} .

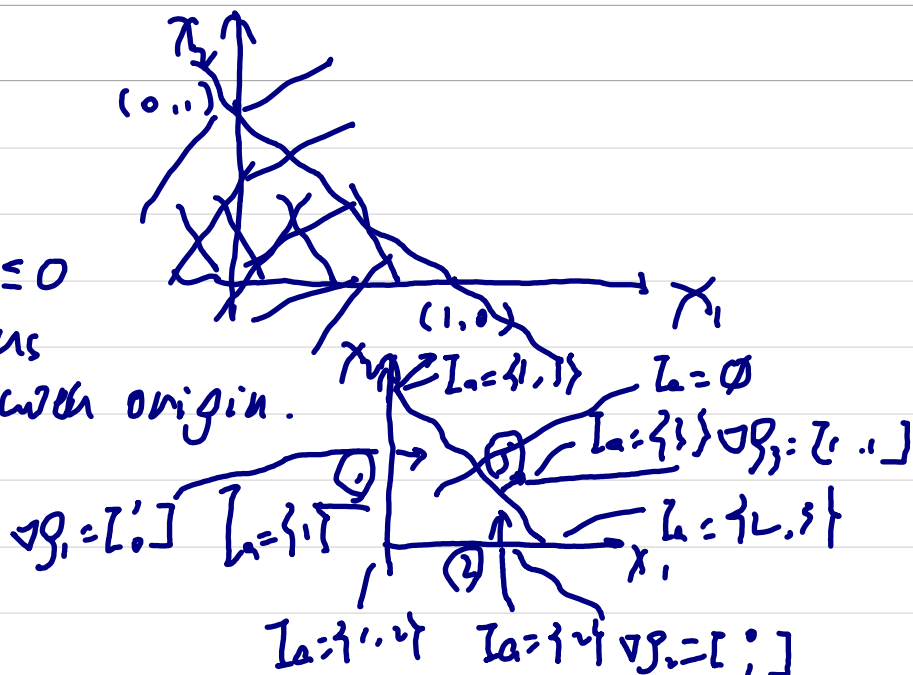
$$x \in (x_1, x_2)$$

$$\begin{array}{l|l} \textcircled{1} & x_1 \geq 0 \\ \textcircled{2} & x_2 \geq 0 \\ \textcircled{3} & x_1 + x_2 = 1 \end{array} \quad \begin{array}{l} -x_1 \leq 0 \\ -x_2 \leq 0 \\ x_1 + x_2 - 1 \leq 0 \end{array}$$

$$(2) \quad \chi_v \neq 0 \quad | \quad -\chi_v = 0$$

$$\textcircled{5} \lambda_1 + \lambda_2 = 1 \mid \lambda_1 + \lambda_2 - 1 \leq 0$$

(ii) decide side with origin.

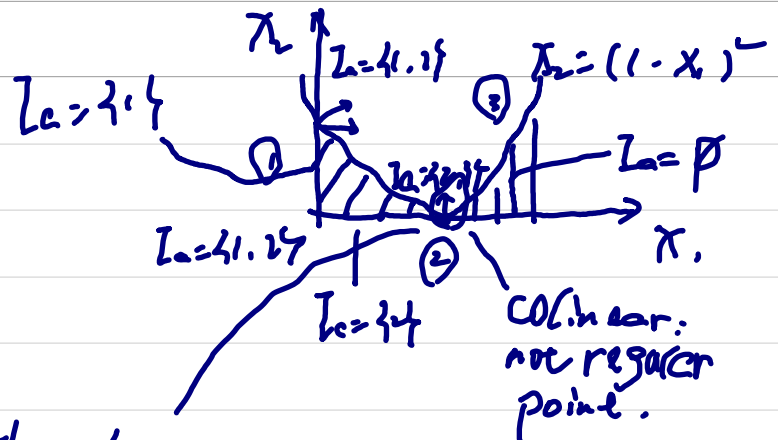


A point \tilde{x} is a regular point for the constraint if the gradients of all active constraint of \tilde{x} are linearly independent.

Given $\tilde{x} \rightarrow \begin{bmatrix} \nabla p_1(\tilde{x}) & \nabla p_2(\tilde{x}) & \dots & \nabla p_m(\tilde{x}) \\ \nabla h_1(\tilde{x}), \dots, \nabla h_l(\tilde{x}) \end{bmatrix}$

All points in the interior of the admission set are regular points.

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ x_2 &\leq (1-x_1)^2 \end{aligned}$$



the co-linear set to be discussed
around the non-regular one by one.
points can change
drastically with small change
of constraints.

$$P_0: \begin{cases} \min f(x) \\ g(x) = 0, x \in \mathbb{R}^n \\ h(x) \leq 0 \end{cases} \quad \text{Define the Lagrangian of the problem}$$

$$L(x, \lambda, p) = f(x) + \lambda' g(x) + \underbrace{p' h(x)}_{\text{penalty}}$$

$$x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m, p \in \mathbb{R}^l$$

First order necessary condition of optimisation

Given P_0 . Suppose that x_0 is a local solution of the problem. Suppose x_0 is a regular point. Then there exist (unique) multipliers λ_0, p_0 such that

$$\begin{aligned} \nabla_x L(x_0, \lambda_0, p_0) &= 0 \\ \nabla_x h(x_0, \lambda_0, p_0) &= g(x_0) = 0, h(x_0) = 0 \\ p_0 &\geq 0, (p_0)' h(x_0) &= 0 \end{aligned} \quad \begin{aligned} \forall i: (p_0)_i h_i(x_0) &= 0 \\ \text{complementary problem} \end{aligned}$$

$$\begin{aligned}
 (y_*)' h(x_*) &= \underbrace{(f_*)' h_1(x_*)}_{\leq 0} + \underbrace{h_1(x_*)}_{\leq 0} \leq 0 \\
 &\underbrace{(f_*)' h_2(x_*)}_{\leq 0} + \dots \\
 &\vdots \\
 &\underbrace{(f_*)' h_r(x_*)}_{\leq 0}
 \end{aligned}$$

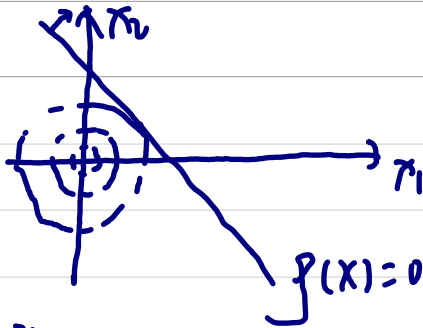
Strict complementary

$$(f_*)' h_1(x_*) = 0 \quad \left[\begin{array}{l} (f_*)' = 0, h_1(x_*) < 0 \\ (f_*)' > 0, h_1(x_*) = 0 \end{array} \right] \\
 (f_*)' = 0, h_1(x_*) = 0$$

At x_*, y_* , for the strict complementary condition hold if $(f_*)' h_1(x_*) = 0$

$$\min \frac{x_1^2 + x_2^2}{2}$$

$$x_1 + x_2 - 1 = 0$$



$$L(x, \lambda) = \frac{x_1^2 + x_2^2}{2} + \lambda(x_1 + x_2 - 1)$$

$$0 = \nabla_x L = \begin{bmatrix} x_1 + \lambda \\ x_2 + \lambda \end{bmatrix} \Rightarrow \begin{aligned} x_1 &= -\lambda \\ x_2 &= -\lambda \end{aligned}$$

$$0 = g(x) = [x_1 + x_2 - 1] \quad -2\lambda - 1 = 0 \Rightarrow \lambda = -\frac{1}{2}$$



CONSTRAINT ELIMINATION
CONSTRAINED \rightarrow UNCONSTRAINED.

$$\min_{x_1} \frac{x_1^2 + (1-x_1)^2}{2} \Rightarrow \min_{x_1} \frac{2x_1^2 - 2x_1 + 1}{2} \sim \tilde{f}$$

$$\nabla \tilde{f} = 4x_1 - 2 = 0 \Rightarrow x_1 = \frac{1}{2}, \quad x_2 = \frac{1}{2}$$

$$\nabla_{xx} L = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

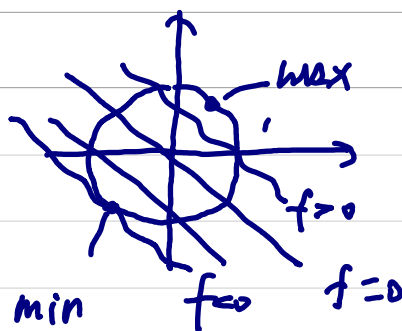
$s^T \nabla_{xx} L s > 0$ for $s \neq 0$ and s such that

$$\begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{bmatrix} s = 0 \quad [1 \ 1] \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} \alpha \neq 0.$$

$$[\alpha \ -\alpha] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} = 2\alpha^2 > 0$$

$$\min x_1 + x_2$$

$$x_1^2 + x_2^2 - 1 \leq 0$$



$$L = (x_1 + x_2) + \rho(x_1^2 + x_2^2 - 1)$$

$$0 = \nabla_x L : \begin{cases} 1 + 2\rho x_1 \\ 1 + 2\rho x_2 \end{cases} \quad \begin{matrix} \rho > 0 \\ x_1^2 + x_2^2 - 1 \leq 0 \end{matrix}$$

Complementary condition

$$\rho(x_1^2 + x_2^2 - 1) = 0$$

$$L \quad \rho = 0 \Rightarrow x_1^2 + x_2^2 - 1 \leq 0$$

$$\begin{cases} \begin{bmatrix} 1+0 \\ 1+0 \end{bmatrix} = 0 \end{cases} \quad \text{X}$$

$$\underline{\rho > 0} \Rightarrow \begin{bmatrix} 1+2\rho x_1 \\ 1+2\rho x_2 \end{bmatrix} = 0 \Rightarrow x_1 = x_2 = -\frac{1}{2\rho}$$

$$\left(\begin{array}{l} x_1^2 + x_2^2 - 1 = 0 \Rightarrow \frac{1}{4\rho^2} \cdot 2 = 1 \Rightarrow \rho_1 = \frac{\sqrt{2}}{2} \\ \rho_2 = -\frac{\sqrt{2}}{2} \end{array} \right) \quad \text{X}$$

$$x_1 = x_2 = -\frac{\sqrt{2}}{2}$$

$$\nabla_{xx} L = \begin{bmatrix} 2\rho & 0 \\ 0 & 2\rho \end{bmatrix} = 2\rho I > 0 \text{ since } \rho > 0$$

$$\frac{dh}{dx} = [2x_1 \ 2x_2] \Big|_{x_\phi} = -\left[\frac{\sqrt{2}}{2} \ \frac{\sqrt{2}}{2}\right] \quad s' \nabla_{xx} L s = [\alpha \ -\alpha] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\frac{dh}{ds} = -\frac{\sqrt{2}}{2} [1 \ 1] \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = 0 \quad s = \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} (\alpha \neq 0)$$

$$\begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} = 4\rho \tilde{x}_0$$

Sufficient conditions

- $L(x, \lambda, p) = f(x) + \lambda' f(x) + p' h(x)$
- Regularity of constraints
- Strict complementary conditions

Given p_0 . Suppose that there exist $\lambda_0, \lambda_0, p_0$ such that

- x_0 is a regular point
 - The strict complementary condition holds
 - $0 = \nabla_x L(\lambda_0, \lambda_0, p_0) \cdot h(x_0) \leq 0$
 $0 = f(x_0) \quad p_0 \geq 0$
 $(p_0)' h(x_0) = 0$
- necessary condition

(+) consider $\underbrace{s' \nabla_{xx} L(x_0, \lambda_0, p_0) s}_{n \times n}$

If this is strictly positive for all $s \neq 0$ and

$$\left[\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial h}{\partial x} \end{bmatrix} (x_0) \right] s = 0$$

Then x_0 is a local strict optimiser.