4.2.1 CONVEX OPTIMIZATION IN STANDARD FORM

$$\begin{cases} \text{minimize} & f_0(x) \\ \text{subject to:} & f_i(x) \leq 0, \quad i=1,\dots,m \\ & a_i^T x = b_i \quad i=1,\dots,p \end{cases}$$

CONVEX OPTIMIZATION PROBLEM IN STANDARD FORM

WHERE $f_0(x)$, $f_i(x)$ ARE CONVEX

COMPARING TO GENERAL CASE:
1) f_0 IS CONVEX
2) f_i ARE CONVEX
3) EQUALITY CONSTRAINTS ARE AFFINE

$\Rightarrow$ FEASIBLE SET IS CONVEX, BECAUSE IT IS INTERSECTION OF CONVEX SETS

RELATED PROBLEMS

1) WHEN f_0 IS QUASICONVEX, THEN THE PROBLEM IS A QUASICONVEX OPTIMIZATION PROBLEM
2) IF f_0 IS CONCAVE AND WE WANT TO MAXIMIZE IT, THEN BY CONSIDERING $-f_0$ WE READILY ARRIVE AT PREVIOUS CASE.

QUESTION: IS THE FOLLOWING PROBLEM A CONVEX OPTIMIZATION PROBLEM?

$$\begin{aligned} \text{minimize} \quad & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} \quad & f_1(x) = \frac{x_1}{1+x_2^2} \leq 0 \\ & h_1(x) = (x_1+x_2)^2 = 0 \end{aligned}$$

NO, BUT IT IS EQUIVALENT TO ONE

QUESTION: IS THIS PROBLEM CONVEX?

minimize $f(x)$ CONVEX

subject to $x \in C$ CONVEX

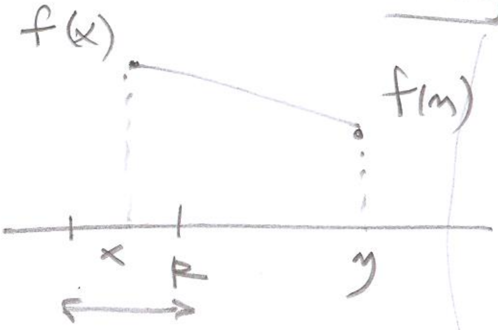
ACCORDING TO OUR NOTATION, NO. WE NEED TO SPECIFY FUNCTIONS, WHICH IS EASY. WE CALL IT AN ABSTRACT FORM CONVEX OPTIMIZATION PROBLEM

4.2.2 LOCAL AND GLOBAL OPTIMA

FUNDAMENTAL PROPERTY: x LOCALLY OPTIMAL $\Rightarrow$ GLOBAL OPTIMUM! PROOF: LET x BE LOCALLY OPTIMAL $\Rightarrow$

$$f_0(x) = \inf \{ f_0(z) \mid z \text{ FEASIBLE}, \|z-x\|_2 \le R \}$$

FOR SOME R. x IS ALSO FEASIBLE. LET x NOT BE GLOBAL OPTIMUM, SO THERE IS y WITH $\|y-x\|_2 > R$ SUCH THAT $f_0(y) < f_0(x)$.



LET $z = (1-\theta)x + \theta y$,

$$\theta = \frac{R}{2\|y-x\|_2}$$

THEN

$$\|z-x\|_2 = \|\theta(y-x)\|_2 = \frac{R}{2\cancel{\|y-x\|_2}} \cancel{\|y-x\|_2}$$

AND

$$f_0(z) \le (1-\theta) f_0(x) + \theta f_0(y) < f_0(x),$$

WHICH IS A CONTRADICTION

PROPERTY IS NOT TRUE FOR QUASICONVEX PROBLEMS

4.2.3 OPTIMALITY CRITERION FOR DIFFERENTIABLE f_0

WE KNOW THAT $\nabla f(x_0) = 0$ IS SUFFICIENT FOR GLOBAL OPTIMALITY BUT IT IS NOT NECESSARY, DUE TO THE EXISTENCE OF CONSTRAINTS. IN FACT, FOLLOWING HOLDS:

PROPERTY: LET f_0 BE DIFFERENTIABLE

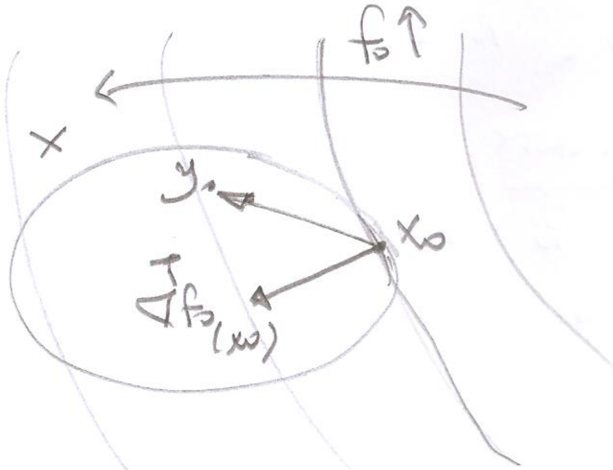
x_0 IS OPTIMAL $\Longleftrightarrow$ x_0 IS FEASIBLE AND

$$(1) \quad \nabla f_0^T(x_0)\,(y - x) \geq 0 \quad \forall y \in X,$$

WHERE X IS FEASIBLE SET:

$$X = \{ x \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p \}$$

INTUITION:



SO IF $\nabla f_0^T(x) \neq 0$, WE ARE ALSO OK, PROVIDED MOST OF X IS IN THE DIRECTION OF $\nabla f_0^T(x)$

ALSO NOTE THAT $-\nabla f_0(x)$ DEFINES A SUPPORTING PLANE AT x_0.

PROOF: WE WILL USE THE KNOWN PROPERTY

$$f_0(y) \geq f_0(x) + \nabla f_0^T(x)\,(y - x) \quad \forall x, y,$$

IF f_0 IS CONVEX.

$(\Rightarrow)$ LET x_0 SATISFY (1). THEN IF $y \in X$, WE HAVE $f_0(y) \geq f_0(x)$ BY ABOVE PROPERTY.

SO CONDITION IS SUFFICIENT.

LET x_0 BE OPTIMAL, AND ASSUME THAT (1) DOES NOT HOLD, SO THERE IS A $y \in X$ SUCH THAT

$$\nabla f_0^T(x_0)(y - x_0) < 0.$$

LET THE POINT $z(t) = ty + (1-t)x$, $t \in [0,1]$.

$z(t)$ IS FEASIBLE, BECAUSE X IS CONVEX. HOWEVER, FOR SOME $t > 0$, $f_0(z(t)) < f_0(x_0)$.

INDEED:
$$\frac{d}{dt} f_0(z(t)) \Big|_{t=0} = \nabla f_0^T(x)(y - x) < 0,$$

AND SO WE HAVE A CONTRADICTION

<u>SPECIAL CASE OF PROPERTY:</u> PROBLEM IS UNCONSTRAINED.

WE HAVE ASSUMED f_0 IS DIFFERENTIABLE, ITS DOMAIN IS OPEN, BY DEFINITION. (THIS IS AN IMPORTANT TECHNICALITY) THE PROPERTY BECOMES:

$$x_0 \text{ IS OPTIMAL} \iff \nabla f(x_0) = 0.$$

<u>PROOF.</u> ($\Rightarrow$) LET x_0 OPTIMAL $\Rightarrow$ $\forall y \in X$,

$\nabla f^T(x_0)(y - x_0) \geq 0$. SINCE X IS OPEN, WE TAKE

$$y = x_0 - t\nabla f(x_0) \text{ FOR SMALL } t > 0,$$

AND
$$\nabla f^T(x_0)(y - x_0) = -t\|\nabla f(x_0)\|_2^2 \geq 0 \Rightarrow$$

$$\|\nabla f(x_0)\|_2 = 0 \Rightarrow \boxed{\nabla f(x_0) = 0}$$

($\Leftarrow$) IF $\nabla f(x_0) = 0$, THEN PREVIOUS PROPERTY AUTOMATICALLY HOLDS

EXAMPLE MINIMIZE $f_0(x) = \frac{1}{2} x^T P x + q^T x + r$.

WITH NO LOSS OF GENERALITY, WE CAN TAKE P SYMMETRIC, I.E. $P^T = P$. INDEED, LET P NOT SYMMETRIC. THEN:

$$\left(\frac{1}{2} x^T P x\right)^T = \frac{1}{2} x^T P^T x,$$ THEREFORE

$$\boxed{\frac{1}{2} x^T P^T x = \frac{1}{2} x^T P x}$$

(THIS IS EXPECTED, BECAUSE

$$x^T P x = \sum_{1 \le i, j \le n} p_{ij} x_i x_j$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_j x_i$$

$$\binom{i \to j}{j \to i} = \sum_{j=1}^{n} \sum_{i=1}^{n} p_{ji} x_i x_j = x^T P^T x \Big)$$

THEREFORE: $x^T P x = \frac{1}{2} x^T P x + \frac{1}{2} x^T P^T x = \frac{1}{2} x^T (P + P^T) x$,

WHERE $P + P^T$ IS SYMMETRIC.

NOW BACK TO THE PROBLEM OF MINIMIZING $f_0(x)$

IF P IS NOT $\succeq 0$, THEN $\inf f_0(x) = -\infty$.

INDEED, $\exists x_0$: $x_0^T P x_0 < 0$. LET

$$g(t) = f_0(t x_0) = \frac{1}{2} t^2 x_0^T P x_0 + (q^T x_0) t + r \to -\infty$$

SO LET $P \succeq 0$, SO THAT $f_0(x)$ IS CONVEX (ITS HESSIAN IS P). THE PROBLEM IS UNCONSTRAINED, SO WE FIND GRADIENT.

<u>DERIVATION:</u> $\nabla\left(\frac{1}{2}x^T P x\right) = Px$.

$$\frac{\partial}{\partial x_i}\left(\frac{1}{2}x^T P x\right) = \frac{\partial}{\partial x_i}\left[\frac{1}{2}p_{ii}x_i^2 + \frac{1}{2}\sum_{j\neq i} 2p_{ij}x_i x_j\right]$$

$$= p_{ii}x_i + \sum_{j\neq i} p_{ij}x_j = \sum_j p_{ij}x_j$$

THEREFORE $\nabla\left(\frac{1}{2}x^T P x\right) = Px$

ALSO, $\nabla q^T x = q$

THEREFORE: $\nabla f_0(x) = 0 \Leftrightarrow \boxed{Px + q = 0}$

WE CONSIDER CASES:

- IF $P \succ 0$, THEN $\det P \neq 0 \Rightarrow \boxed{x_0 = -P^{-1}q}$ IS UNIQUE MINIMIZER.
- IF P HAS 0 EIGENVALUES, THEN TWO CASES:
  1) $q \notin \mathcal{R}(P) \Rightarrow$ NO SOLUTION

     INTUITION: THERE IS A DIRECTION ALONG WHICH $\frac{1}{2}x^T P x$ IS ZERO, AND ONLY THE LINEAR TERM HAS AN EFFECT $\Rightarrow$ $\inf = -\infty$
  2) $q \in \mathcal{R}(P) \Rightarrow$ INFINITE SOLUTIONS. THE LINEAR TERM IS ZERO ALONG DIRECTION IN WHICH THE QUADRATIC FORM IS ALSO ZERO

EXAMPLE:
ANALYTIC CENTERING

$$\text{minimize} \quad f_0(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x),$$

$$\text{subject to} \quad Ax < b,$$

$$\text{WITH} \quad A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}$$

THE FUNCTION IS DIFFERENTIABLE AND THE FEASIBLE SET IS OPEN, SO NECESSARY AND SUFFICIENT CONDITION IS:

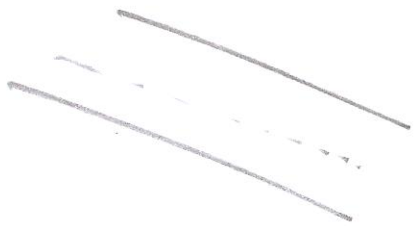
$$\nabla f_0(x) = 0 \iff \nabla \sum_{i=1}^{m} \log(b_i - a_i^T x) = 0 \iff$$

$$\sum_{i=1}^{m} \nabla \log(b_i - a_i^T x) = 0 \iff \boxed{\sum_{i=0}^{m} \frac{a_i}{b_i - a_i^T x} = 0} \quad (A)$$

THE FOLLOWING CASES EXIST (PROOF IS EXERCISE)

1) NO SOLUTION TO (A) $\iff$ f_0 UNBOUNDED BELOW

2) MANY SOLUTIONS. FOR EXAMPLE:



3) UNIQUE SOLUTION $\iff$

OPEN POLYHEDRON $\{x \mid Ax < b\}$ IS NON-EMPTY AND BOUNDED

EQUALITY CONSTRAINTS ONLY

CONSIDER THE PROBLEM $\quad$ minimize $f_0(x)$

subject to: $Ax = b$

WHAT DOES THE OPTIMALITY CONDITION BECOME?

INTUITION: THE FUNCTION SHOULD INCREASE IN A DIRECTION THAT IS ORTHOGONAL TO THE AFFINE SET $Ax = b$.

OPTIMALITY CONDITION FOR x_0: $\quad \nabla f_0^T(x_0)\,(y - x_0) \geq 0$

$\forall y$ SUCH THAT $\quad Ay = b$.

LET THE FEASIBLE SET $\quad y = x_0 + v, \quad v \in N(A)$

WE NEED $\quad \nabla f_0^T(x_0)\,(x_0 + v - x_0) \geq 0 \iff \nabla f_0^T(x_0)\, v \geq 0$

$\forall v \in N(A)$. BUT $N(A)$ IS A SPACE, SO WE HAVE $\quad \nabla f_0^T(x_0)\, v = 0 \quad \forall v \in N(A) \iff$

$$\nabla f_0^T(x_0) \perp N(A)$$

{ FUNDAMENTAL PROPERTY OF LINEAR ALGEBRA:

$$N(A)^{\perp} = R(A^T)$$ }

THEREFORE, $\quad \nabla f_0(x_0) \in R(A^T) \iff$

$$\exists v \in \mathbb{R}^p : \quad \nabla f_0(x) + A^T v = 0$$

$\nabla f_0(x)$ IS A LINEAR COMBINATION OF THE ROWS OF A

LAGRANGE MULTIPLIER OPTIMALITY CONDITION!
WE WILL SEE IT AGAIN

ASIDE:
PROOF THAT $N(A)^{\perp} = R(A^T)$

$v \in N(A) \Leftrightarrow Av = 0 \Leftrightarrow \forall w, \; (Av)^T w = 0 \Leftrightarrow$

$\forall w, \; v^T(A^T w) = 0 \Leftrightarrow v \in [R(A^T)]^{\perp}$

SECOND ASIDE: INTUITION FOR $\nabla f_0(x_0) + A^T \nu = 0$.

IF I AM IN x_0, I CAN MOVE IN ANY DIRECTION t AND GO TO $x_0 + t$ AS LONG AS $A(x_0+t) = b \Leftrightarrow$

$At = 0 \Leftrightarrow \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} t = 0 \Leftrightarrow a_1 t = 0, \; a_2 t = 0, \ldots \; a_p t = 0$

$\Leftrightarrow$ t^T IS ORTHOGONAL TO ALL ROWS OF a_p

SO THE CONDITION SAYS THAT THE GRADIENT OF f_0 SHOULD BE IN THE LINEAR SPACE SPANNED BY THEM.

MINIMIZATION OVER THE NONNEGATIVE ORTHANT

LET US CONSIDER THE SPECIAL CASE:

minimize $f_0(x)$
subject to $x \succeq 0$

THE OPTIMALITY CONDITION IS

$$x \succeq 0, \quad \nabla f_0^T(x)(y - x) \geq 0 \quad \forall y \succeq 0$$

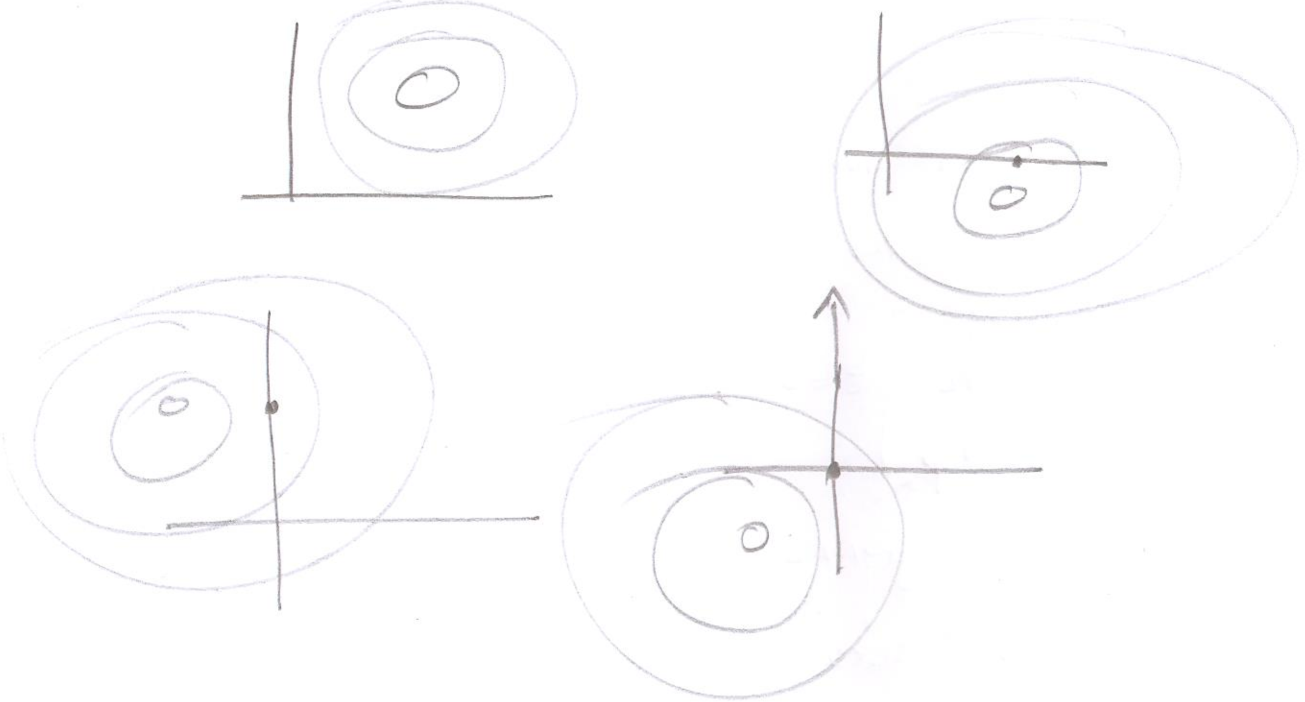
BECAUSE y CAN HAVE ARBITRARILY COMPONENTS, WE NEED $\nabla f_0^T \geq 0$. BUT THE CONDITION MUST HOLD FOR $y = 0$, SO WE ALSO NEED $\nabla f_0^T(x)\, x \leq 0$ AND WE ALSO HAVE $x \succeq 0$ THEREFORE, WE NEED

$x_i \, (\nabla f_0(x))_i = 0 \quad i = 1, \ldots, n$, AND SUMMARIZING,

$$x \geq 0, \quad \nabla f_0(x) \geq 0, \quad x_i(\nabla f_0(x))_i = 0$$

COMPLEMENTARITY CONDITION.

INTUITION (FOR 2D)



4.2.4 EQUIVALENT CONVEX PROBLEMS

1. ELIMINATING EQUALITY CONSTRAINTS

$$\begin{cases} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0 \\ & Ax = b. \end{cases}$$

WE HAVE SEEN THAT WE ELIMINATE EQUALITY CONSTRAINTS BY SETTING

$x = x_0 + Fz$, WITH $\mathcal{R}(F) = \mathcal{N}(A)$

THE PROBLEM BECOMES

$$\begin{aligned} &\text{minimize} && f_0(x_0 + Fz) \\ &\text{subject to} && f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{aligned}$$

WHICH IS ALSO CONVEX.

SO, IN PRINCIPLE, WE COULD HAVE DEFINED THE CONVEX PROBLEM AS A PROBLEM WITH NO EQUALITY, BUT KEEPING THE EQUALITIES CAN HELP

2. INTRODUCING EQUALITY CONSTRAINTS

WE CAN ALSO DO THE INVERSE. IF WE HAVE

$f_i(A_i x + b)$, WE CAN INTRODUCE $y_i = A_i x + b$, ($A_i \in \mathbb{R}^{k_i \times n}$) AND AGAIN WE HAVE A CONVEX PROBLEM.

3. SLACK VARIABLES

$f_i(x) \leq 0$ BECOMES $f_i(x) + s = 0$. SO, WE NEED INEQUALITIES TO BE LINEAR! OTHERWISE, NEW PROBLEM IS NOT CONVEX!

4. EPIGRAPH PROBLEM FORM

THE EPIGRAPH FORM IS

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && f_0(x) - t \leq 0 \\ & && f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & && a_i^T x = b_i. \end{aligned}$$

THE NEW CONSTRAINT $f_0(x) - t \leq 0$ IS CONVEX, AS CONVEX COMBINATION OF CONVEX FUNCTIONS, SO NEW PROBLEM IS CONVEX.

THEREFORE, ALL CONVEX PROBLEMS CAN BE CONVERTED TO CONVEX PROBLEMS WITH LINEAR CONSTRAINTS!

5. MINIMIZING OVER SOME VARIABLES

$$\left\{ \begin{aligned} &\text{minimize} && f_0(x_1, x_2) \\ &\text{subject to} && f_i(x_1) \leq 0, \quad i = 1, \ldots, m_1 \\ & && \tilde{f}_i(x_2) \leq 0, \quad i = 1, \ldots, m_2 \end{aligned} \right\}$$

CONSIDER THE FOLLOWING PROBLEM WHICH IS ASSUMED CONVEX

DEFINE $\tilde{f}_0(x_1) = \inf \{ f_0(x_1, z) \mid \tilde{f}_i(z) \leq 0, \ i = 0, \ldots, m_2 \}$

THE PROBLEM IS EQUIVALENT TO:

$$\left\{\begin{array}{ll} \text{minimize:} & \tilde{f}_0(x_1) \\ \text{subject to:} & f_i(x_i) \leq 0, \quad i=1,\dots,m_2 \end{array}\right\}$$

IS THE NEW PROBLEM CONVEX?

WE DO A DIVERSION TO CONVEX FUNCTIONS:

3.2.5 MINIMIZATION

PROPERTY: LET $f(x,y)$ CONVEX IN (x,y), AND C CONVEX NONEMPTY SET. THEN $g(x) = \inf_{y \in C} (f(x,y))$

WHERE ITS DOMAIN IS $\operatorname{dom} g = \{x \mid (x,y) \in \operatorname{dom} f$ for some $y \in C\}$,

I.E. THE PROJECTION ON THE x COORDINATES) IS CONVEX.

Proof: LET x_1, x_2.

$$g(\theta x_1 + (1-\theta)x_2) = \inf_{y \in C} \{ f(\theta x_1 + (1-\theta)x_2, y) \} \quad (*)$$

LET $\epsilon > 0$.

WE SELECT y_1, y_2 SUCH THAT
$$f(x_1, y_1) \leq g(x_1) + \epsilon$$
$$f(x_2, y_2) \leq g(x_2) + \epsilon.$$

THEN
$$\begin{aligned} (*) &\leq f(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \\ &\leq \theta f(x_1, y_1) + (1-\theta) f(x_2, y_2) \\ &\leq \theta g(x_1) + \theta\epsilon + (1-\theta)\epsilon + (1-\theta) g(x_2) \\ &= \theta g(x_1) + (1-\theta) g(x_2) + \epsilon. \end{aligned}$$

THEREFORE:

$$g(\theta x_1 + (1-\theta)x_2) \leq \theta g(x_1) + (1-\theta) g(x_2) + \epsilon \qquad \forall \epsilon > 0$$

$$\Rightarrow \quad g(\theta x_1 + (1-\theta)x_2) \leq \theta g(x_1) + (1-\theta) g(x_2)$$

EXAMPLE:
$$\text{dist}(x, S) = \inf_{y \in S} \|x - y\|$$

IS THE DISTANCE BETWEEN A POINT $x \in \mathbb{R}^n$ AND A SET.

IF S IS CONVEX, THEN $\text{dist}(x, S)$ IS CONVEX FUNCTION OF x.

EXAMPLE: LET $h(y)$ CONVEX, $y \in \mathbb{R}^k$. LET.
$$g(x) = \inf\{h(y) \mid Ay = x\} = \inf_{y \in \mathbb{R}^k}\{f(x,y)\}$$

WHERE
$$f(x,y) = \begin{cases} h(y) & \text{IF } Ay = x \\ \infty & \text{OTHERWISE} \end{cases}$$

$f(x,y)$ IS CONVEX, SO $g(x)$ IS CONVEX.

— · —

RETURNING TO THE ORIGINAL PROBLEM OF MINIMIZATION OVER SOME VARIABLES, $\tilde{f}_0$ IS CONVEX! SO NEW PROBLEM IS ALSO CONVEX

4.2.5 QUASICONVEX OPTIMIZATION

$$\begin{aligned} &\text{minimize} && f_0(x) && \leftarrow \text{QUASICONVEX} \\ &\text{subject to} && f_i(x) \le 0, \quad i, \dots, m && \leftarrow \text{CONVEX} \\ & && Ax = b \end{aligned}$$

(IF QUASICONVEX, WE CAN SUBSTITUTE WITH CONVEX)

IT CAN HAVE LOCAL OPTIMA WHICH ARE NOT GLOBAL OPTIMA!

RECALL FIRST-ORDER CONDITION FOR DIFFERENTIABLE f:

f QUASICONVEX $\Leftrightarrow$ $\text{dom} f$ IS CONVEX AND $\forall x, y \in \text{dom} f$
$$f(y) \le f(x) \Rightarrow \nabla f(x)^T (y - x) \le 0$$

PROPERTY x IS OPTIMAL IF IT BELONGS TO FEASIBLE SET X AND $\nabla f_0^T(x)(y-x) > 0 \quad \forall y \in X$

FOLLOWS IMMEDIATELY FROM PREVIOUS PROPERTY.

COMPARE WITH CONDITION FOR CONVEX FUNCTIONS: LET $x \in X$

1) x OPTIMAL $\Leftrightarrow$ $\nabla f_0^T(x)(y-x) \geq 0 \quad \forall y \in X$

THEREFORE:
1) NEW CONDITION IS SUFFICIENT, NOT NECESSARY
2) IT NEEDS $\nabla f_0^T(x) \neq 0$

SOLVING QUASICONVEX OPTIMIZATION PROBLEMS

BASIC IDEA: WE CAN CREATE A FAMILY OF CONVEX FUNCTIONS $\varphi_t : \mathbb{R}^n \to \mathbb{R}$ SUCH THAT

$$f_0(x) \leq t \Leftrightarrow \varphi_t(x) \leq 0$$

WE CAN ALWAYS FIND SUCH A FAMILY. AT THE VERY LEAST, THE FOLLOWING WILL WORK:

$$\varphi_t(x) = \begin{cases} 0, & f_0(x) \leq t, \\ \infty, & \text{OTHERWISE} \end{cases} \qquad \varphi_t(x) = \text{dist}\left(x, \{z \mid f_0(z) \leq t\}\right)$$

LET p^* BE THE OPTIMAL VALUE, THEN $p^* \leq t$ IFF THE FEASIBILITY PROBLEM IS FEASIBLE:

$$\begin{aligned} &\text{FIND} \quad x \\ &\text{SUBJECT TO} \quad \varphi_t(x) \leq 0 \quad \leftarrow \text{NEW} \\ &\qquad f_i(x) \leq 0, \quad i = 0, \dots, m \\ &\qquad Ax = b \end{aligned}$$

THEREFORE WE CAN SOLVE THE ORIGINAL PROBLEM USING A SEQUENCE OF CONVEX PROBLEMS TOGETHER WITH THE BISECTION METHOD