In summary:

Inequality constrained problem:

$$\min_{x\in\mathbb{R}^n}\{V(x) : h(x)\le 0\} = \hat{V} \qquad \text{(ICP)}$$

$\boxed{h:\mathbb{R}^n \to \mathbb{R}^m}$

Perturbation function:

$$\pi(y) = \min_{x\in\mathbb{R}^n}\{V(x) : h(x)\le y\}$$

$$\pi(0) = \hat{V}$$

Dual function:

$$d(\lambda) = \min_{y\in\mathbb{R}^m}\{\pi(y) + \lambda^T y\} \qquad (18.1)$$

Easily-evaluatable formula for $d(\lambda)$:

$d(\lambda)$ is also given by:

$$d(\lambda) = \min_{x\in\mathbb{R}^n}\{V(x) + \lambda^T h(x)\} \qquad (18.2)$$

only when $\lambda \geqslant 0$ ← $\boxed{\text{meaning: } \lambda_i \geqslant 0 \ \forall i=1,..m}$

Using definition (18.1) of $d(\lambda)$ until further notice, an outline of duality theory for (ICP) follows:

Exactly as in Th (14.2):

$$d(\lambda) \le \hat{V} \quad , \forall \lambda \in \mathbb{R}^m$$

$$\Downarrow$$

(18.3) $\quad d(\lambda) \le \hat{V} \quad , \forall \lambda \geqslant 0$

Assuming from now on that:

- π is convex on $\mathbb{R}^m$
- π is differentiable at $y=0$

we have:

(18.4) $\quad d(\hat{\lambda}) = \hat{V}$ where $\hat{\lambda} = -\nabla\pi(0) \geqslant 0$

$\boxed{\text{unknown}}$

Hence, from (18.3) and (18.4):

(18.5) $$\underbrace{\max_{\lambda \geq 0} d(\lambda)}_{\boxed{\text{the dual problem for (ICP)}}} = \hat{V} = \underbrace{\min_{x \in \mathbb{R}^n} \{V(x) : h(x) \leq 0\}}_{\boxed{\text{the primal problem (ICP)}}}$$

Since $\lambda \geq 0$ in the dual, we can use the easily-evaluatable formula for d in the dual, making the solution of the dual problem potentially practicable.

We can find $\hat{x}$ for (ICP) from $\hat{\lambda}$:

actually unique, found by maximizing d.

(18.5) Th Under the above conditions:

$$\arg\min (\text{ICP})$$
$$\|$$
$$\left\{\arg\min_{x \in \mathbb{R}^n} \{V(x) + \hat{\lambda}^T h(x)\}\right\} \cap \{x \in \mathbb{R}^n : h(x) \leq 0\}$$
$$\|$$
$$\left\{\arg\min_{x \in \mathbb{R}^n} \{V(x) + \hat{\lambda}^T h(x)\}\right\} \cap \{x \in \mathbb{R}^n : \hat{\lambda}^T h(x) = 0\}$$

from: $\nabla \pi(0)^T h(\hat{x}) = 0$ Th (17.6)

and $\hat{\lambda} = -\nabla \pi(0)$

If the singleton assumption is valid:

$$\underset{\boxed{\text{for (ICP)}}}{\hat{x}} = \tilde{x} = \arg\min_{x \in \mathbb{R}^n} \{V(x) + \hat{\lambda}^T h(x)\}.$$

One can go on to use augmented duals to derive a multiplier alg. for (ICP), all similar to equality constrained problem.

Quadratic programming

(very important for various applications)

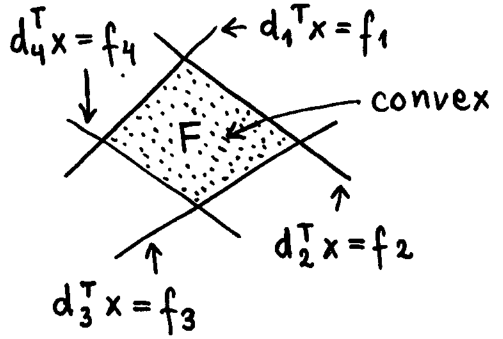
Means of solving of a quadratic program, which is an optimization problem of the kind:

$$\min_{x \in \mathbb{R}^n} \{\text{standard quadratic} : \underbrace{D}_{\in \mathbb{R}^{m \times n}} x \leq f\}$$

Practical applications of QP's include:

- finding search directions for some advanced algs.
- constrained optimal control problems

Feasible set $\quad F = \{x \in \mathbb{R}^n : D x \leq f\}$, $\quad D = \begin{bmatrix} d_1^T \\ \vdots \\ d_m^T \end{bmatrix}$



The simple geometry of such F enables special solution methods to be developed for Q.P.s.

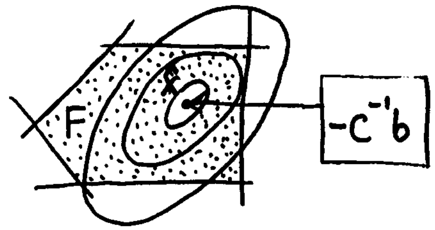
Solution of Q.P.s

For our standard quadratic, the unconstrained global minimizer is $-C^{-1}b$.

For the QP: there are 2 possible situations:

I $\quad -C^{-1}b \in F$

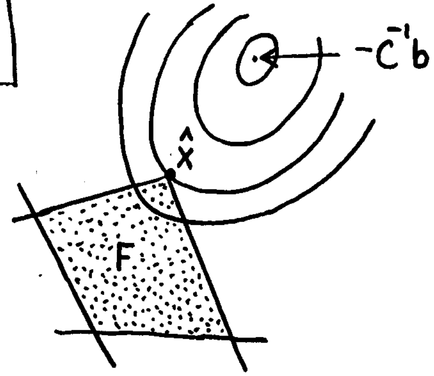
i.e. $D(-C^{-1}b) \leq f$



Obviously in this case $\hat{x} = -C^{-1}b$ ... $\boxed{\text{Q.P. solved}}$

II $\quad -C^{-1}b \notin F$ $\leftarrow \boxed{\text{more likely}}$

i.e. $D(-C^{-1}b) \nleq F$



How to find $\hat{x}$?

... Computationally

In case II :

clearly always $\hat{x} \in \underbrace{\text{boundary of } F}$

$\left\{ x \in \mathbb{R}^n : \text{some selection of the rows of } D^T x = f \text{ are satisfied as equality constraints and } D^T x \leq f \right\}$

Hence <u>$\hat{x}$ (for Q.P.) can be found by</u> :

- for each possible selection of the rows of $D^T x = f$, find the corresponding optimal x

  easy as V = quadratic and each row of $D^T x = f$ is a linear equality constraint

- check for feasibility of any such x

- choose the feasible x which gives the least optimization cost (least among all optimal xs in F).

  $\Rightarrow$ optimal $\hat{x}$ (*)

<u>Duality for Q.Ps</u>

The primal Q.P. is :

$$\min_{x \in \mathbb{R}^n} \{ \underbrace{a + b^T x + \tfrac{1}{2} x^T C x}_{\boxed{V(x) = \text{convex}}} : \underbrace{Dx - f \leq 0}_{\boxed{h(x) \text{ all } h_i(x) \text{ are convex}}} \}$$

$\Downarrow$

primal problem is convex

$\Downarrow$ (17.3)

Π = convex

(*) many Q.P. algs work like this but choose the xs to be tried in an intelligent way so that not all of them are needed.

$$d(\lambda) = \min_{x \in \mathbb{R}^n} \{ V(x) + \lambda^T h(x) \}$$

$$= \min_{x \in \mathbb{R}^n} \{ (a - \lambda^T f) + (b + D^T\lambda)^T x + \tfrac{1}{2} x^T C x \}$$

$$\underset{\text{minimum of a standard unconstrained quadratic}}{=} (a - \lambda^T f) - \tfrac{1}{2} (b + D^T\lambda)^T C^{-1} (b + D^T\lambda)$$

Fact: the duality requirement that:

$\pi(y)$ be differentiable at $y = 0$

can be replaced by the <u>Slater constraint qualification</u>:

$$\exists\, x \in \mathbb{R}^n : \ Dx < f$$

$\boxed{\text{i.e. } (Dx)_i < f_i,\ \forall i}$

a very reasonable condition,
<u>assumed to be satisfied here</u>

Provided the Slater constraint qualification is satisfied:

$$\underbrace{\min_{x \in \mathbb{R}^n} \{ \underbrace{V(x)}_{\text{quadratic}} : Dx \leq f \}}_{\text{Q.P. primal}} = \hat{V} \underset{(18.5)}{=} \underbrace{\max_{\substack{\lambda \geq 0 \\ \lambda \in \mathbb{R}^m}} \underbrace{d(\lambda)}_{\text{quadratic in } \lambda}}_{\text{the dual Q.P.}}$$

$\Downarrow$

$\hat{x}$ for Q.P. can be found easily from
$\hat{\lambda}$ for the dual Q.P. $\leftarrow$ (18.6)

Often : $m \ll n$:

then constrained optimization of the m λ_is can be much easier than direct constrained optimization of the n x_is,

$\Rightarrow$ hence the potential computational significance of the dual problem

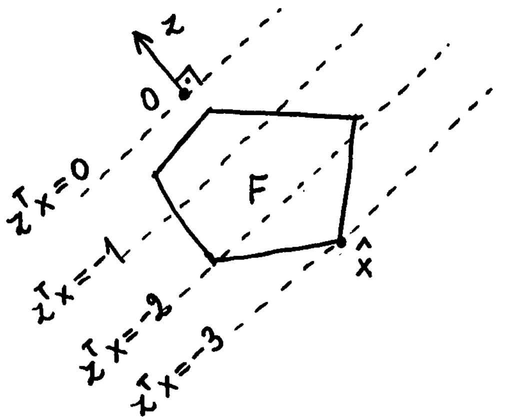
Linear programming

(also very important for applications)

A typical linear program is:

$$\min_{x \in \mathbb{R}^n} \{ z^T x : Dx \le f \}$$

Applications:
- optimization in economics
- optimal control



The corners of F are called its extreme points.

(18.7) **Th** A global minimizer for a L.P. is an extreme point of F.

$\{ x : Dx \le f \}$

Consequence:

Many algorithms for solving LPs find all extreme points of F, then find an extreme point giving least cost, and that is an $\hat{x}$. Actually, those algs. iterate so as to consider only extreme points giving lower cost than does the current extreme point.

Duality for L.P.:

The primal problem is:

$$\min \{ \underbrace{z^T x}_{\boxed{V(x) = \text{convex}}} : \underbrace{Dx - f}_{\boxed{h(x),\ \text{with convex } h_i}} \le 0 \}$$

$\Downarrow$

primal problem = convex

$\Downarrow$

π = convex

$$d(\lambda) = \min_{x \in \mathbb{R}^n} \{ V(x) + \lambda^T h(x) \}$$

$$= \min_{x \in \mathbb{R}^n} \{ -\lambda^T f + (z + D^T \lambda)^T x \}$$

(18.8) **Lemma** $d(\lambda) = -\infty$ if $z + D^T\lambda \neq 0$

Proof: If $z + D^T\lambda \neq 0$, consider

$$x \triangleq -\theta(z + D^T\lambda) \quad , \theta\text{ - scalar}$$

Then: $-\lambda^T f + (z + D^T\lambda)^T x = -\lambda^T f - \theta \| z + D^T\lambda \|^2$

which can be made as negative as desired by making the scalar θ big enough.

Hence one can make $-\lambda^T f + (z + D^T\lambda)^T x$ as negative as desired. by choosing x appropriately, which is what is meant by:

$$\underbrace{\min_{x \in \mathbb{R}^n} \left\{ -\lambda^T f + (z + D^T\lambda)^T x \right\}}_{d(\lambda)} = -\infty$$

QED.

Now: assume the Slater constraint qualification is valid.

Then:

$$\min_{x \in \mathbb{R}^n} \{ z^T x : Dx \leq f \} \leftarrow \boxed{\text{the primal LP.}}$$

$$\| $$

$$\hat{V} = \max_{\lambda \geq 0} d(\lambda)$$

$$= \max_{\substack{\lambda \geq 0 \\ z + D^T\lambda = 0}} d(\lambda) = \underbrace{\max_{\substack{\lambda \geq 0 \\ z + D^T\lambda = 0}} (-\lambda^T f)}$$

since $d(\lambda) = -\infty$ which cannot be maximal if $z + D^T\lambda \neq 0$ we can filter out such non-maximizing λs by adding the constraint $z + D^T\lambda = 0$

the dual L.P. potentially useful when $m \ll n$

Note: you could not hope to easily write a good alg. for LP or QP. ... use the excellent (and very sophisticated) library programs.