DISCRETE OPTIMIZATION

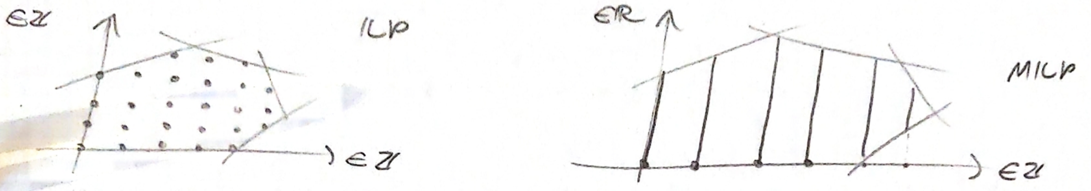
3-1 IP MODELS

Def a mixed integer linear programming (MILP) problem is

$$\begin{aligned} \min\ & \underline{c}^T \underline{x} \\ \text{st}\ & A\underline{x} \geq \underline{b} \\ & \underline{x} \in \underbrace{\mathbb{Z}^{n_1}}_{\substack{\text{integer part} \\ (n_1 \text{ vars in } \mathbb{Z})}} \times \underbrace{\mathbb{R}^{n_2}}_{\substack{\text{mixed part} \\ (n_2 \text{ vars in } \mathbb{R})}} \end{aligned} \qquad \left(\begin{aligned} & A \in \mathbb{Z}^{m \times (n_1+n_2)} \\ & \underline{c} \in \mathbb{Z}^{n_1+n_2} \\ & \underline{b} \in \mathbb{Z}^{m} \end{aligned} \right)$$

If - $x_j \in \mathbb{Z}$ $\forall j$ then we just call it ILP (integer linear prog.)
- $x_j \in \{0,1\}$ $\forall j$ we talk of BILP (or 0/1 ILP, for binary lin. prog.)

BILPs are NP-hard, and (M)ILP are at least as difficult. We have $n = n_1+n_2$ variables and m constraints. Feasible regions examples are



MODELING TECHNIQUES AND EXAMPLES

(1) a binary variable allows to model a choice yes/no between two alternatives.

EX Binary knapsack problem

variables $x_i = \begin{cases} 1 & \text{if object } i \text{ is chosen} \\ 0 & \text{otherwise} \end{cases}$

model $\max \sum p_i x_i$
st $\sum w_i x_i \leq b$ (capacity)
$x_i \in \{0,1\}$ $\forall i$

(2) Binary variables allow to model the association between two entities. The decide with x_{ij} if to associate object/group i to group/object j.

(3) forcing constraints (or linking-variables constr.) to impose that a decision x can be made only if other y was made.

EX Uncapacitated facility location (UFL)

variables x_{ij} : fraction of demand of client i served by depot j;

$y_j = \begin{cases} 1 & \text{if depot } j \text{ is opened} \\ 0 & \text{otherwise} \end{cases}$

Model $\min \sum\sum \tilde{c}_{ij} x_{ij} + \sum f_j y_j$ — these become c_{ij} - we can reformulation factor out the demand of constr (1, the demand)

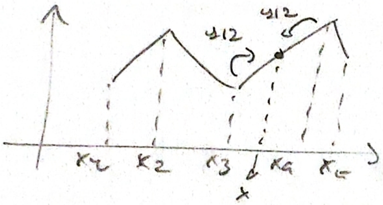
st $\sum_j x_{ij} = 1$ $\forall i$ (satisfy clients demands)

$\sum_i x_{ij} \leq M \cdot y_j$ $\forall j$ OR $x_{ij} \leq y_j$ $\forall i \forall j$

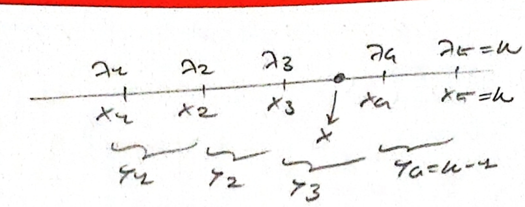
(these constr are to impose that $y_j = 0 \Rightarrow x_{ij} = 0$, $\exists x_{ij} > 0 \Rightarrow y_j = 1$)

$0 \leq x_{ij} \leq 1$
$y_j \in \{0,1\}$ $\forall j$

(4) Piecewise linear cost functions: we have more nodes and a piecewise function there defined, and we want to either describe all the xs in that interval



$\rightsquigarrow$



we can define and then $y_i = \begin{cases} 1 & \text{if } x \text{ is inside interval } i \text{, i.e. } (x_i, x_{i+1}) \\ 0 & \text{otherwise} \end{cases}$

$\min f(x)$
st $x \in [x_1, x_k]$

$\Longleftrightarrow$ (can be reformulated as)

$\min \sum_{i=1}^{k} \lambda_i f(x^i)$

st $\sum \lambda_i = 1$ (convex comb)
$\sum y_i = 1$ (only one interval)
$\lambda_i \le y_{i-1} + y_i \quad \forall i = 2, \dots, k-1$
$\lambda_1 \le y_1$
$\lambda_k \le y_{k-1}$

$\lambda_i \ge 0$
$y_i \in \{0,1\} \ \forall i$

(5) modelling with exponentially many constraints, eg on graphs we use still assignments binary variables but the we work on subsets of edges/vertices to build constr.

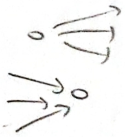
<u>EX</u> asymmetric travelling salesman problem ($\boxed{\text{ATSP}}$)

<u>variables</u> $x_{ij} = \begin{cases} 1 & \text{if we select arc } (i,j) \\ 0 & \text{otherwise} \end{cases}$

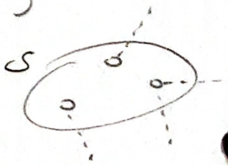
<u>Model</u> $\min \sum_{(i,j) \in A} x_{ij} c_{ij}$

st $\sum_{(i,j) \in \delta^+(i)} x_{ij} = 1 \quad \forall i$ (one outgoing arc)

$\sum_{(i,j) \in \delta^-(i)} x_{ij} = 1 \quad \forall i$ (one incoming arc)



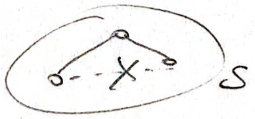
$\sum_{(i,j) \in \delta^+(S)} x_{ij} \ge 1 \quad \forall S \subset V, S \neq \emptyset$ (cutset inequalities)



OR

$\sum_{(i,j) \in E(S)} x_{ij} \le |S| - 1 \quad \forall S \subseteq V, \ 2 \le |S| \le n-1$

(sub-tour elimination)



$x_{ij} \in \{0,1\} \ \forall i \forall j$

$\left(\begin{array}{l} \delta^+(S) = \{(i,j) \in A : i \in S, j \notin S\} \\ E(S) = \{(i,j) \in A : i \in S, j \in S\} \end{array} \right)$

(6) With binary variables we can also impose disjunctive constraints, i.e. let just one (among more) be valid

$\underline{a}_1 \cdot \underline{x} \le b_1$
$\underline{a}_2 \cdot \underline{x} \le b_2$ $\rightsquigarrow$ we assign them binary variables y_1, y_2 to decide which to activate

$\Rightarrow$ transform them into $\underline{a}_i \cdot \underline{x} - b_i \le \underbrace{M(1 - y_i)}$
together with $y_1 + y_2 = 1$

≤ 0 if $y_i = 1$
$\le M$ if $y_i = 0$, and $M \gg 0$ to make that constraint trivially always true

(7) Linearization of product variables.

- product of binary variables

  $z = y_1 \cdot y_2$ with both $y_i \in \{0,1\}$

  $\Rightarrow$ we introduce z and add the following constr on it:

  $$z \le y_1$$
  $$z \le y_2$$
  $$z \ge y_1 + y_2 - 1$$

- product of binary and bounded continuous variable

  $z = y \cdot x$ with $y \in \{0,1\}$ and $x \in [0,u]$

  $\Rightarrow$ we introduce z and the constr:

  $$z \le y \cdot u$$
  $$z \le x$$
  $$z \ge 0$$
  $$z \ge x - (1-y)u$$

  $\begin{cases} z \le 0 \ (z \ge 0) & y = 0 \\ z \ge x & y = 1 \end{cases}$

3.2 STRONG FORMULATIONS

For MILP, the formulation of the problem is crucial. Often we would try to simplify them, being easier to solve, by relaxing the integrality constr.
So now we talk about relaxations.

MILP	LP-relaxation
$z_{MILP} = \min c^T x$	$z_{LP} = \min c^T x$
s.t. $Ax \ge b$	s.t. $Ax \ge b$
$x \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$	$x \in \mathbb{R}^{n_1+n_2}$

Prop. We know that the LP relaxation problem will find in general a better sol (i.e. for min problem, $z_{LP} \le z_{MILP}$).
If that sol however is integer then is optimal also for the MILP problem.

Def. a polyhedron $P = \{x \in \mathbb{R}^{n_1+n_2} : Ax \ge b, x \ge 0\} \subset \mathbb{R}^{n_1+n_2}$ is a formulation for a mixed integer set $X \subseteq \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$ if we know that $X = P \cap (\mathbb{Z}^{n_1} \times \mathbb{R}^{n_2})$.

formulations are relaxed, continuous

Working with X a (possibly mixed) integer set, there could be different polyhedrons that lead to the same X, i.e. different formulations. We look for stronger ones.

Def. Given X a mixed integer set $X \subseteq \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$, and two formulations P_1 and P_2 for it, we say that

P_1 is stronger than P_2 $\iff$ $P_1 \subset P_2$

$\iff$ P_1 is "smaller" (is more similar to X) than P_2

To show that a formulation P_1 is stronger than another one P_2 we need to show that:

- $P_1 \subset P_2$ (i.e. all the elts of P_1 are also elts of P_2)
- $\exists$ a elt $x \in P_2$ but $x \notin P_1$

Remember that with formulations we are dealing with relaxed problems.

Thm (Meyer). Let $X \subseteq \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$ be the mixed integer feasible set of a MILP with rational coeffs. Then conv(X) is a rational polyhedron, and all its extreme points belong to X.

Def. So we say that a polyhedron $P \subseteq \mathbb{R}^{n_1+n_2}$ is the ideal formulation of X if conv(X) $= P$.

Ideal since we could solve the MILP problem by solving its relaxation of P, i.e. solving a simple (easy) LP.

The idea will be to derive the strong formulations, by reasoning about which constraints to add or how to model the problem

EX Perfect matching problem (PM) on a graph with $|V|$ even

Variables $x_e = \begin{cases} 1 & \text{if we select edge } e \\ 0 & \text{otherwise} \end{cases}$

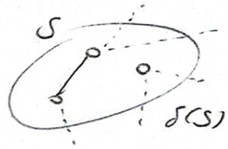
Model

$$\min \sum_{e \in E} c_e x_e$$

$$\text{st} \sum_{e \in \delta(i)} x_e = 1 \quad \forall i \in V \quad \text{(only one out-coming edge)}$$

$$\left[\sum_{e \in \delta(S)} x_e \geq 1 \quad \forall S \subset V \text{ with } |S| \text{ odd} \right]$$

this added constraint is needed to reach an ideal formulation



$$x_e \in \{0,1\} \quad \forall e \in E$$

Or we can go to extended formulations, which is when we include additional variables, still to reach stronger formulations

EX Uncapacitated lot sizing (ULS)
Determine the production plan for the next n time periods.

- original formulation:
  - x_t = amount produced at time t
  - y_t = 1 if production occurs at time t
  - s_t = amount in stock at the end of time t
- extended formulation:
  - w_{ut} = amount produced in time u to satisfy the demand of future period t

(from this we can derive x_t, y_t and s_t, w's is an extension)

How do we compare the strength of extended formulations?

Def. Given a polyhedron $P \subset \mathbb{R}^n \times \mathbb{R}^{n'}$, the orthogonal projection onto the x-subspace $\mathbb{R}^n$ is the polyhedron

$$\text{proj}_{\underline{x}}(P) = \{\underline{x} \in \mathbb{R}^n : \exists \underbrace{\underline{w} \in \mathbb{R}^{n'}}_{\text{variables of the extended formulation}} \text{ st } (\underline{x}, \underline{w}) \in P\}$$

We can obtain the projection using the Fourier-Motzkin method, which for simple systems is just the idea of

- isolating a variable x_i in the equations
- remove it but add all the inequalities resulting from lower and upper bounds of x_i, on the other variables

Def. a compact formulation is a formulation with a number of variables/constraints polynomial in the instance size.

However as usual we like strong and ideal formulations, which are often achieved introducing variables and/or constraints.
So often compact formulations are weak.

3.3 EASY ILP PROBLEMS (TU)

Consider a generic ILP problem and its relaxation:

$$\begin{array}{ll} \min & c^T x \\ \text{st} & Ax = b \\ & x \in \mathbb{Z}^n_+ \end{array} \quad (\text{ILP}) \qquad \Bigg| \qquad \begin{array}{ll} \min & c^T x \\ \text{st} & Ax = b \\ & x \geq 0 \end{array} \quad \left(\begin{array}{c}\text{LP} \\ \text{relax}\end{array}\right)$$

Assume that $A \in \mathbb{Z}^{m \times n}$, with $n \geq m$ (more variables than constr) and also $b \in \mathbb{Z}^m$. We know that if the LP relaxation has an optimal sol which is integer, then is optimal also for the ILP.

- any LP has the optimal sol on a vertex
- to each vertex corresponds (at least) one basic feasible sol, ie

$$x = \left(\frac{x_B}{x_N}\right) = \left(\frac{B^{-1} b}{0}\right) \qquad A = \left(\boxed{B} \,\vdots\, N \right)$$

B is a basis of A, ie a $m \times m$ non-singular submatrix

$\Rightarrow$ $Ax = b$ becomes
$Bx_B + Nx_N = b$
$Bx_B = (b - Nx_N)$

- if an optimal basis B has det of ± 1, then the sol $x = (x_B, x_N)$ will be integer and then optimal for ILP.

How do we guarantee that last condition? With this

Def. A matrix $A \in \mathbb{Z}^{m \times n}$ is totally unimodular (TU) if every squared submatrix B of A has determinant $-1, 0,$ or 1.

Prop. We have the following properties:
(P1) A is TU $\Leftrightarrow$ A^T is TU
(P2) A is TU $\Leftrightarrow$ $(A : I_m)$ is TU
(P3) A' obtained permuting or changing the sign of some cols/rows of A is TU $\Leftrightarrow$ A was TU

Why this condition is useful?

Thm. If $A \in \mathbb{Z}^{m \times n}$ is TU, b is integral ($\in \mathbb{Z}^m$) and
$P(b) = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \neq \emptyset$ relaxation of $Ax = b$
OR
$P(b) = \{x \in \mathbb{R}^n : Ax \geq b, x \geq 0\} \neq \emptyset$ relaxation of $Ax \geq b$

$\Rightarrow$ then all vertices of $P(b)$ are integral
$\Rightarrow$ we to optimally solve the ILP it suffices to solve its LP relaxation

In practice, how do we check (after that we built the model for the problem) that a matrix is TU? We use the properties P1-P3 and the following

Prop (a sufficient condition for TUness)

(1) $a_{ij} \in \{-1, 0, 1\}$ $\forall i \, \forall j$
(2) each col of A contains at most two non-zero coefficients
(3) we can divide the rows into two groups I_1 and I_2 st for all cols j which have two non-zero coeffs we have that

$$\sum_{i \in I_1} a_{ij} = \sum_{i \in I_2} a_{ij}$$

$\Rightarrow$ A is TU

eg for this A

$$\begin{array}{ccccc} 0 & 0 & -1 & 0 & \\ 0 & 0 & 1 & 0 & I_1 \\ 1 & 1 & 0 & 0 & \\ -1 & 0 & 0 & 1 & \\ \hline 0 & 0 & 0 & 1 & I_2 \\ \downarrow & \downarrow & \downarrow & \downarrow & \\ \text{check } 0=0 & \text{dont check} & \text{check } 0=0 & \text{check } 1=1 & \end{array}$$

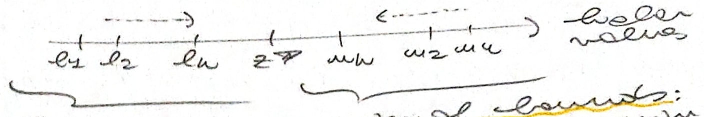
The TUness is useful since it gives us ideal formulations.

But for some problems there are better polynomial-time solvers/algos that exploit the problem structure.

3.4 RELAXATIONS AND BOUNDS

In general algorithms solving discrete optimization problems (like ILPs) provide as result a sequence of upper bounds and a sequence of lower bounds.

$z^* = \min c(\underline{x})$
st $\underline{x} \in X$



Dual bounds:
- are LBs for a min pb
- are obtained through a relaxation

Primal bounds:
- are UBs for a min pb
- are obtained through heuristics (to get only feasible sol $\underline{x} \in X$)

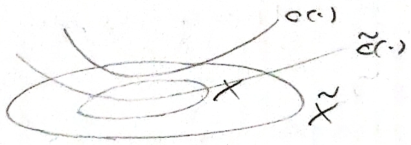
Where the termination criterion will be $(u_k - l_k) \le \varepsilon$ (to have a guaranteed quality on the obtained sol.

RELAXATIONS

Def. Given the problems: we say that

(P) $z^* = \min\{c(\underline{x}) : \underline{x} \in X \subseteq \mathbb{R}^n\}$

(RP) $\tilde{z} = \min\{\tilde{c}(\underline{x}) : \underline{x} \in \tilde{X} \subseteq \mathbb{R}^n\}$



RP is a relaxation of P $\Leftrightarrow$
- $X \subseteq \tilde{X}$ (bigger feasible region)
- $\tilde{c}(x) \le c(\underline{x})\ \forall x \in X$ (smaller obj function on X)

Prop. If RP is a relaxation of P then $\tilde{z} \le z^*$ (we get a better optimal sol)

Prop. Let $\tilde{\underline{x}}_{RP}$ be the optimal sol of the RP.
- if $\tilde{\underline{x}}_{RP} \in X$ (is feasible for P)
- and $\tilde{c}(\tilde{\underline{x}}_{RP}) = c(\tilde{\underline{x}}_{RP})$ (obj func equal) → $\tilde{\underline{x}}_{RP}$ is also optimal for P

There are different possible relaxation methods:

(1) the LP (linear programming) relaxation: the only one we knew till now. Just we remove the integrality constr.

(2) relaxation by elimination: simply remove one or more constraints. Is a very weak relaxation idea.

(3) surrogate relaxation (SR): replace a subset of constr with their linear combination with multipliers $\lambda_i \ge 0$. Then look for the dual problem to get the best bound.

(4) lagrangian relaxation (LR): remove a "difficult" constr and add a term in the obj function that penalizes its violation.

EX Multiple binary knapsack $x_{ij} = \begin{cases} 1 & \text{if we take object } j \text{ into knapsack } i \\ 0 & \text{otherwise} \end{cases}$

$\max \sum_i \sum_j p_j x_{ij}$

st (1) $\sum_j w_j x_{ij} \le W_i$ ∀ knapsack i

(2) $\sum_i x_{ij} \le 1$ ∀ item j

(3) $x_{ij} \in \{0,1\}$

- SR: max //
  st (2), (3)
  $$\sum_i \lambda_i \left(\sum_j w_j x_{ij}\right) \le \sum_i \lambda_i (W_i)$$

- LR: $\max\ // + \sum_j \underbrace{u_j}_{\ge 0} \underbrace{\left(1 - \sum_i x_{ij}\right)}_{<0 \text{ (so penalization) if constr is violated}}$
  st (1)
  (3)

(5) combinatorial relaxations: when we try to simplify the problem by solving it recalling when possible easier structures to deal with (like spanning trees on graphs).

Prop. SR and LR dominate the relaxation by elimination
Prop. SR dominates LR.
But in practice LR is more used since it leads to easier problems to solve.

HEURISTICS

(1) Greedy methods: build a feasible sol piece-by-piece, selecting the option that leads to the best local improv. and without reconsidering past choices.

(2) Local search methods: we iteratively try to improve a current feasible sol.

- start from initial x_0 feasible
- at iteration k:
    - find a best sol x' in $N(x_k)$, the neighborhood of x_k (set of nearby feasible sols)
    - if $c(x') < c(x_k)$, i.e. we improved, then continue with the search, otherwise return x_k

(3) Metaheuristics: try to escape from local optima. An example is Tabu search, where we allow moves to nearby sets even if they worsen the objective funct, and we store the moves/directions taken to avoid come back.

EX
- for (1): in the binary knapsack problem we can order the items based on profit/weight ratio
- for (2): in STSP we can remove two random non-adjacent edges e_1 and e_2 and replace them to get another path
- for (3): in UFL we can for any set S of candidate depots to be opened consider $S \cup \{i\}$ or $S \setminus \{j\}$ as nearby sets in which to move

3.6 CUTTING PLANE METHODS

We know that we strive for strong formulations. But getting directly to an ideal one is complex, so the idea is rather to "refine" an initial formulation by adding ineqs.

Def. $\pi \cdot x \le \pi_0$ is a valid inequality for $X \subseteq \mathbb{R}^n$ ⟺ $\pi \cdot x \le \pi_0 \ \forall x \in X$
⟺ i.e. if it is satisfied by all points of X

How do we use them? we could
(1) add them a priori, but this will make the problem difficult to solve even for the LP relaxation or from a branch & bound approach, due to the huge # of constr.

(2) generate them when needed, and this is the cutting plane idea.

actually here is "for a given $x' \notin X$", so that we cut it away)

Consider a generic ILP $\min\{c^T x : x \in X = P \cap \mathbb{Z}^n\}$.
Def. a cutting plane is a candidate ineq $\pi \cdot x \le \pi_0$ st
- it is valid $\forall x \in X$, i.e. $\pi \cdot x \le \pi_0 \ \forall x \in X$
- it is not valid outside of X, i.e. $\pi \cdot x' > \pi_0 \ \forall x' \in P$ (and $\notin X$)

(6 In this way there is no need to get to the ideal formulation necessarily, but instead we just "trim" out regions of P bringing integer vertices out as optimal sols.

Method: initialize $P' = P = \{\underline{x} \in \mathbb{R}^n_+ : A\underline{x} \leq \underline{b}\}$, i.e. the initial LP relaxation relaxation. Then:

(1) Solve the current LP relaxation $\min\{\underline{c}^T\underline{x} : \underline{x} \in P'\}$. Let $\underline{x}^*_{LP}$ be the optimal sol.

(2) If $\underline{x}^*_{LP} \in \mathbb{Z}^n \Rightarrow$ end, it is also optimal for ILP

Else:

(2a) Solve the separation problem: given $\underline{x}^*_{LP} \notin X$, a family of ineq, find a ineq that separates $\underline{x}^*_{LP}$ from $conv(X)$ (or establish that there is none of them).

If ineq found $\Rightarrow$ that is the cutting plane and we update $P' = P \cap \{\underline{x} \in \mathbb{R}^n_+ : \underline{\pi} \cdot \underline{x} \leq \pi_0\}$ and go back to (1)

Else: stop, we can't further improve

We have different examples/methods for generating cutting planes.

- Chvátal-Gomory procedure: generate ineqs via linear combinations of the constraints and rounding.
  - Let $X = P \cap \mathbb{Z}^n = \{\underline{x} \in \mathbb{R}^n_+ : A\underline{x} \leq \underline{b}\} \cap \mathbb{Z}^n$ so that $X = \{\underline{x} \in \mathbb{Z}^n_+ : \sum_{j=1}^n \underbrace{\underline{A}_{\bullet j}}_{j\text{-th col of } A} x_j \leq \underline{b}\}$
  - Choose the multipliers vector $\underline{u} \in \mathbb{R}^m_+$ (not nec $\sum u_i = 1$) and consider $\sum_j (\underline{u}^T\underline{A}_{\bullet j}) x_j \leq \underline{u}^T\underline{b}$
  - Since $\lfloor \underline{u}^T\underline{A}_{\bullet j} \rfloor \leq \underline{u}^T\underline{A}_{\bullet j}$ and $x_j \geq 0$ we get that $\sum_j \underbrace{\lfloor \underline{u}^T\underline{A}_{\bullet j} \rfloor}_{int} \underbrace{x_j}_{int} \leq \underline{u}^T\underline{b}$ is a ineq for P, X, and $conv(X)$.
  - Then we can get to the stronger version rounding both sides, and it is a cutting plane since is valid for X and $conv(X)$ (but not nec for P).

$$\sum_{j=1}^{m} \lfloor \underline{u}^T\underline{A}_{\bullet j} \rfloor x_j \leq \lfloor \underline{u}^T\underline{b} \rfloor$$

This method is very strong since we have that Theorem (Chvátal). any ineq for any X can be obtained by applying this procedure a finite number of times. Also, given any fractional $\underline{x}^*_{LP}$, there exist a multiplier vector $\underline{u}$ st the CG procedure cuts it from P.

Def. Call first Chvátal closure of P the set $P_1 = \{\underline{x} \in \mathbb{R}^n_+ : A\underline{x} \leq \underline{b}, A^1\underline{x} \leq \underline{b}^1\}$, where $A^1\underline{x} \leq \underline{b}^1$ are all the inequalities obtained by varying the vector $\underline{u} \in \mathbb{R}^m_+$ in the CG procedure

Def. We call Chvátal rank of $conv(X)$ the smallest integer k such that $P_k = conv(X)$.

- Gomory fractional/integer cutting planes: if $\underline{x}^*_{LP}$ is fractional, let x^*_a be a fractional basic variable. Then the integer form of the Gomory fractional cut, generated from a row t of the LP relaxation, is

$$x_a + \sum_{j \in N} \lfloor \bar{a}_{tj} \rfloor x_j \leq \lfloor \bar{b}_t \rfloor$$

non-basic variables

$\bar{a}$ and $\bar{b}$ come from $\underline{x}_B = \underbrace{B^{-1}\underline{b}}_{\bar{b}} + \underbrace{B^{-1}N}_{\bar{A}}\underline{x}_N$

↑ ILP

↓ MILP

- Mixed integer rounding (MIR) inequality, considering now MILPs. If $X = \{(x,y) \in \mathbb{Z} \times \mathbb{R}^+ : x - y \leq b\}$ is, with b not integer,

$$x - \frac{1}{1-\{b\}} y \leq \lfloor b \rfloor$$

, where $\{b\} = b - \lfloor b \rfloor \geq 0$ is the fractional part of b

- Gomory mixed integer (GMI) inequality, still on MILPs, is now we deal with (x^*_{LP}, y^*_{LP}) an optimal basic feasible set of the LP relaxation. We get working as before on $\bar{A}$ and $\bar{b}$ this inequality for a component h st $x^*_h \notin \mathbb{Z}$:

$$x_h + \sum_{\underbrace{j\in N_1}_{\text{non-basic integer vars}}}\left(\lfloor \bar{a}_{tj}\rfloor + \frac{(\{\bar{a}_{tj}\}-\{\bar{b}_t\})^+}{1-\{\bar{b}_t\}}\right)x_j \le \lfloor \bar{b}_t\rfloor + \sum_{\underbrace{j\in N_2}_{\text{non-basic continuous vars}}} \frac{(\bar{a}_{tj})^-}{1-\{\bar{b}_t\}}\,y_j$$

3.7 STRONG VIs FOR STRUCTURED ILPs

Studying the problem structure we can derive strong VIs.

Def. For any $P=\{x\in\mathbb{R}^n_+ : Ax\le b\}$, a VI $\pi\cdot x\le\pi_0$ dominates another VI $\gamma\cdot x\le\gamma_0$ (valid both for P) if

$$\exists u>0 : u\gamma\le\pi \text{ and } u\gamma_0\ge\pi_0 \quad \left(\text{with } \binom{\pi}{\pi_0}\ne u\binom{\gamma}{\gamma_0}\right)$$

which says that the feasible region of the first VI is smaller than the one of the dominated second one.

Def. a VI $\pi\cdot x\le\pi_0$ is redundant in the description of P if exist other VIs for P st their linear combination is valid and dominate that one. ie

$\exists k\ge 2$ VIs $\pi_i\cdot x\le\pi_i$ valid for P, and $\exists u_i>0$ $\forall i=1,\dots,k$: $\left(\sum_{i=1}^{k} u_i\pi_i\right)\cdot x\le\left(\sum_{i=1}^{k} u_i\pi_i\right)$

this new VI dominates it

FACES AND FACETS

The idea then will be to look for VIs which are "necessary" for describing the polyhedron $P=\{x\in\mathbb{R}^n : Ax\le b\}$.

Def. (1) k points $x_1,\dots,x_k\in\mathbb{R}^n$ are affinely independent $\iff$ the $k-1$ vectors they can form are linearly $\perp$ (x_i-x_1 $\forall i=2\to k$)

(2) $\dim(P)=\left(\begin{matrix}\text{max \# of affinely } \perp\\ \text{points of } P\end{matrix}\right)-1$

(3) P is full dimensional $\iff$ $\dim(P)=n$, ie no inequality $a^Tx\le b$ is satisfied at equality by all points $x\in P$

Thm. $\dim(P)=n$ $\implies$ P admits a unique minimal description $P=\{x\in\mathbb{R}^n : a_i^Tx\le b_i \ \forall i=1,\dots,m\}$ where ie each inequality is unique and necessary (their deletion will lead to a different P)

and these necessary inequalities will be the facets ones.

Def. Let $F=\{x\in P : \pi\cdot x=\pi_0\}$ for any VI for P. Then F is a face of P, and we say that the VI represents/defines F. If moreover F is a face and $\dim(F)=\dim(P)-1$, then we call it a facet.

Thm. If P is full dimensional, ie $\dim(P)=n$

a VI is necessary to describe P $\iff$ it defines a facet for P

$\iff$ $\exists\, n$ affinely independent points $\in P$ satisfying it at equality (ie that $\dim(F)=n-1=\dim(P)-1$

Often the P we would be interested in finding its facets will be $\text{conv}(X)$. Then it there is another method to show that a VI defines a facet.

So let $X \subset \mathbb{Z}^n_+$ bounded and $\pi \cdot x \le \pi_0$ a VI for X. Assume conv(X) is bounded and dim(conv(X)) $= n$. To check that the VI is a facet defining one we could:

- apply the definition (or the thm before) finding n points $x_1, \dots, x_n \in X$ satisfying the VI at equality and showing that they are affinely independent
- using and indirect approach:
  (1) select $t(\ge n)$ points $x_1, \dots, x_t \in X$ satisfying the VI at equality. Suppose they all belong to the same hyperplane $\mu^T x = \mu_0$
  (2) identify the hyperplane params $(\mu_1, \dots, \mu_n; \mu_0)$ solving the system
  $$\sum_{j=1}^{n} \mu_j x_j^{(k)} = \mu^T x_k = \mu_0 \qquad \forall k = 1, \dots, t$$
  (3) if the only sol is $\left(\frac{\mu}{\mu_0}\right) = \lambda\left(\frac{\pi}{\pi_0}\right)$ for $\lambda \neq 0$, then the VI is facet defining

BINARY KNAPSACK CASE

We have $X = \{ x \in \{0,1\}^n : \sum_{j=1}^{n} a_j x_j \le b \}$, for $N = \{1, \dots, n\}$.

Def. We call a subset $C \subseteq N$ a cover for X if $\sum_{j \in C} a_j > b$. a cover is minimal if $\forall j \in C$, $C \setminus \{j\}$ is no more a cover.

Covers are useful since provide easily some VIs:

Prop. If $C \subseteq N$ is a cover for X, then the following is a VI:

$$\underbrace{\sum_{j \in C} x_j \le |C| - 1}_{\text{cover inequality}}$$

since surely we have to remove at least 1 item from C to get a maybe feasible sol

But we have more:

Prop.

C is a minimal cover for X $\iff$ that VI is a facet of $P_C = \text{conv}(X) \cap \{ x \in \mathbb{R}^n : x_j = 0 \ \forall j \notin C \}$

ie if we just look for sols only in C (so we ignore items out of C)

(1) Separation problem. If we get (eg by LP relaxation) a fractional sol x^*, how can we find a VI to cut it (ie violated by x^*)?

The idea is to rewrite the VI from $\sum_{j \in C} x_j \le |C| - 1$ to $\sum_{j \in C} (1 - x_j) \ge 1$, and we ask if $\exists$ a cover $\tilde{C}$ that violates that constr. So the idea is to

- solve $\zeta = \min \sum_{j \in N} (1 - x_j^*) z_j$
  st $\sum_{j \in N} a_j z_j > b$ (we still get a cover for X)
  $z \in \{0,1\}^n$
  not nec to opt
- look at the objective function value:
  if $\zeta \ge 1$ then x^* satisfies all cover inequalities and there is no VI that cuts it.

  if $\zeta < 1$ (with sol z^*) then we can add the classical VI using $\tilde{C} = \{ j : z_j^* = 1 \}$ and it cuts x^*.

(2) Strengthening cover inequalities. We have firstly a simple

Prop. If C is a cover for X, the extended cover inequality

$$\sum_{j \in E(C)} x_j \le |C| - 1$$

where $E(C) = C \cup \{ j \in N : a_j \ge a_i \ \forall i \in C \}$ is still valid for X.

The lifting procedure is the systematic way to strengthen a cover inequality. It is useful since

Prop. If C is a minimal cover and $a_j \le b$ $\forall j \in N$, then the lifting procedure will (surely) yield a facet-defining inequality for conv(X).

How does it work?

(1) we have $X = \{x \in \{0,1\}^n : \sum_{j \in N} a_j x_j \le b\}$

(2) get a (possibly) minimal cover $C \subseteq N$ and the corresponding ~~[illegible]~~

(3) set an order for the $j \in N \setminus C$

(4) consider x_i, $i \in N \setminus C$ and find the largest α_i st

$$\alpha_i x_i + \sum_{j \in C} x_j \le |C| - 1$$

i.e. considering the relevant case of when $x_i = 1$:

$$\alpha_i \le |C| - 1 - \sum_{j \in C} x_j \le$$

$$\le |C| - 1 - \left\{ \begin{array}{l} \max \sum_{j \in C} x_j \\ \text{st } \sum_{j \in C} a_j x_j \le b - \underset{\sim}{a_i} \end{array} \right\}$$

as we set $x_i = 1$ so we occupy the knapsack with its weight

(5) now consider the row as updated with $\sum_{j \in C} x_j + \alpha_i x_i \le |C| - 1$ and repeat from step (4) with the next variable

STSP CASE (AND ATSP)

Ex. Symmetric TSP problem. Like the ATSP but now we have an undirected (→ edges) graph.

Variable $x_e = 1$ if we choose edge e

Model $\min \sum_{e \in E} c_e x_e$

st $\sum_{e \in \delta(v)} x_e = 2$ $\forall v \in V$ (1)

$\sum_{e \in E(S)} x_e \le |S| - 1$ $\forall S \subset V : S \ne \emptyset$ (2)

$x_e \in \{0,1\}$ $\forall e \in E$ (3)

Prop. $\forall S \subseteq V : 2 \le |S| \le n/2$ (and $n \ge 4$) we have that constr (2) defines a facet of conv(X).

For the rest see the slides. Well, normally the idea is to move to the ATSP (which had constr of type (1) splitted in two), and solve its LP relaxation without considering constr (2) of there, but using it to add iteratively cutting planes.

And we can find such cutting plane (to remove a vect x^*_P) by solving a sequence of instances of the min cut problem.

EQUIVALENCE BETWEEN SEPARATION AND OPTIMIZATION

Let $P \subseteq \mathbb{R}^n$ be a (bounded) rational polytope.

Opt problem: given P and $c \in \mathbb{Q}^n$ → find $x^* \in P$ maximizing $c^T x$ over P or establish that P is empty

Sep problem: given P and $x^* \in \mathbb{Q}^n$ → find a cut that separates x^* from P or establish that $x^* \in P$

Thm. the sep problem can be solved in polynomial time in n and len(c) $\iff$ the opt problem can be solved in polynomial time in n and len(c)

So even if a problem (opt) seems hard to solve (as it may have an exp number of constr), actually can be easy / polynomial, if the corresponding sep problem (of the cutting plane approach) is easy to solve

3.8 BRANCH AND CUT

Idea: embed strong valid inequalities (on the problem solution) to speed up the classical branch and bound framework.

I.e. if after solving the LP relaxation we don't get an integer sol, we don't directly branch (on the fractional variables) but we add cuts to the formulation.

Advantages:

(1) strengthening the formulation we get better dual bounds

(2) slow convergence of pure cutting plane methods is mitigated by the branching steps

3.9 COLUMN GENERATION METHOD

Used for ILP problems with an exponential # of variables.

The idea is:

- to enumerate all partially feasible sols
- represent the (additional) constrs as a set packing/covering/partitioning type of formulation
- only generate new variables when needed

It's like we did in the cutting planes method, where we had an exp # of constraints, which were not considered all explicitly but instead included/generated on the fly.

EX Cutting stock problem.

A company produces rolls of width W, but b_i small rolls of width w_i instead are needed. So we cut the large ones according to patterns.

- Kantorovich model:

$$z^{K}_{ILP} = \min \sum_{k \in K} y_k$$

$$\text{st} \sum_{k \in K} x_{ik} \geq b_i \quad \forall i \in I \quad \left(\text{demand constr}\right)$$

$$\sum_{i \in I} w_i x_{ik} \leq W \cdot y_k \quad \forall k \in K \quad \left(\text{width \& var link}\right)$$

$$x_{ik} = \left(\text{\# of times a } i\text{-th small roll is cut in the } k\text{-th large roll}\right) \in \mathbb{Z}^+$$

$$y_k = \left(\text{if we cut the } k\text{-th large roll}\right) \in \{0,1\}$$

- Gilmore and Gomory model

$$A = \left(\cdots \underbrace{\overset{\sim}{\begin{pmatrix}1\\0\\2\\2\end{pmatrix}}}_{j\text{-th pattern}} \cdots \right) \text{amount of rolls in large}$$

cut 1 of width w_1, 0 of w_2, 2 of w_3, ecc

$$z^{GG}_{ILP} = \min \sum_{j \in J} x_j$$

$$\text{st} \sum_{j \in J} a_{ij} x_j \geq b_i \quad \forall i \in I \quad \left(\text{demand constr}\right)$$

$$x_j = \left(\text{\# large rolls cut according to pattern } j\right) \in \mathbb{Z}^+$$

$$j \in J = \left(\text{index set of the patterns}\right)$$

The matrix A is $m \times n$, where n will grow exponentially wrt to m (ie with the number of small rolls $w_1, \dots, w_m$).

$$A = \underbrace{\left(\qquad\qquad\qquad \right)}_{n} \Big\} m$$

A will be a wide matrix

But: - at optimality at most m of the n variables (the x_j) will have non zero values

- rounding optimal LP sols will lead to satisfactory results like

Column generation method:

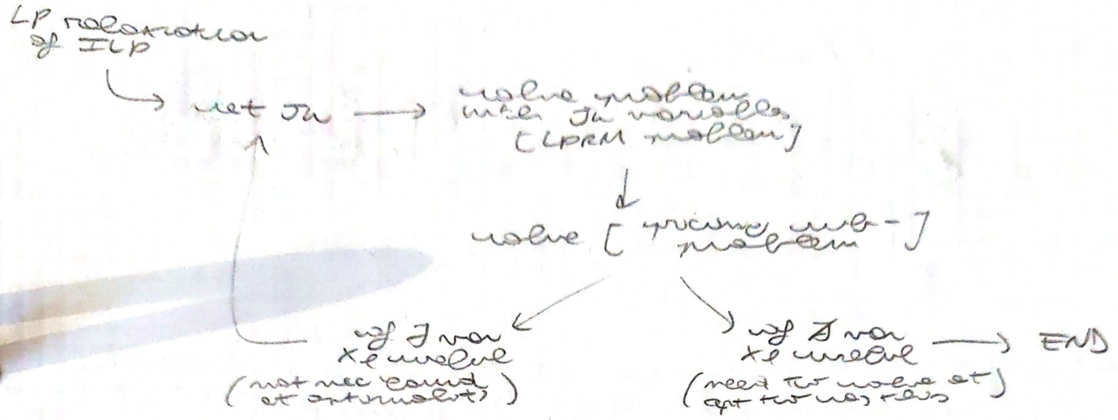
(1) consider the LP relaxation of the ILP, setting $k=0$ and $J_0 \subseteq I$ as the initial subset of variables

(2) solve it, i.e. the LP restricted master (LPRM) problem on the variables selected by J_k

(3a) consider its dual

(3b) solve the pricing subproblem for LPRM with J_k to search for an improving non-basic variable x_e

(a) if $\exists$ such x_e, update $J_k = J_k \cup \{x_e\}$ and go to (2)
otherwise we can't further improve, and LPRM opt sol is also optimal for the overall LP relaxation of the ILP


LP relaxation of ILP
set J_k
solve problem with J_k variables [LPRM problem]
solve [pricing sub-problem]
if $\exists$ var x_e useful (not nec found at optimality)
if $\nexists$ var x_e useful (need to solve at opt to say this)
END

EX Cutting stock problem ($I = \{1, \dots, m\}$) I is the set of the small rolls

$$\left.\begin{aligned} \min\ & \textstyle\sum_{j \in J_k} x_j \\ \text{st}\ & \textstyle\sum_{j \in J_k} a_{ij} x_j \geq b_i \quad \forall i \in I \quad \overset{\text{dual}}{\to y_i} \\ & x_j \geq 0 \quad \forall j \in J_k \end{aligned}\right\} \text{LPRM}$$

dual

$$\left.\begin{aligned} \max\ & \textstyle\sum_{i \in I} y_i b_i \\ \text{st}\ & \textstyle\sum_{i \in I} a_{ij} y_i \leq 1 \quad \forall j \in J_k \\ & y_i \geq 0 \quad \forall i \in I \end{aligned}\right\} \text{LPRM dual}$$

For the pricing subproblem we ask if there is an other pattern (i.e. variable) which is feasible and useful to add to J_k:

(Before writing it, solve the above ones getting x^* and y^*, then:)

$$\left.\begin{aligned} \min\ & \bar{c} = 1 - \textstyle\sum_{i \in I} y_i^* a_i \\ \text{st}\ & \textstyle\sum_{i \in I} w_i a_i \leq W \quad (\text{pattern is feasible}) \\ & a_i \in \mathbb{Z}_+ \end{aligned}\right\} \text{pricing sub-problem}$$

we are searching for a new useful variable to be added

- if $\bar{c}^*$ (opt sol of that) ≥ 0, then there is no useful variable to add
- if $\exists\ \bar{c}^* < 0$ (not nec the opt sol) then we add the variable represented by the a_i's to the set J_k

Observations:
- choice of the initial set of variables/cols J_0 has a strong impact
- for the pricing sub-problem we can use heuristics as we don't have to necessarily solve it at optimality)
- CG method is very flexible and used in practice
- CG can be also included in a B&B (→ branch and price framework).

3.10 LAGRANGIAN RELAXATION

Consider a generic ILP in the form

$$\min\{\underline{c}^T\underline{x} : \underbrace{D\underline{x} \geq \underline{d}}_{\text{complicating constraints}},\ \underbrace{A\underline{x}\geq \underline{b},\ \underline{x}\in\mathbb{Z}^n}_{\text{otherwise this contained in just having } \underline{x}\in X\subseteq\mathbb{R}^n}\}$$

Idea of Lagrangian relaxation: remove those complex constr but add, for each of them, a penalizing term (for their violation) in the objective function.

$$z^* = \min \underline{c}^T\underline{x} \quad \text{st } D\underline{x} \geq \underline{d},\ \underline{x}\in X\subseteq\mathbb{R}^n \qquad (2) \text{ primal problem}$$

here we are at $\underline{u}$ fixed

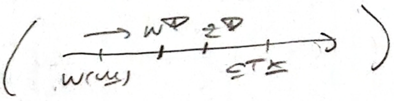
$$w(\underline{u}) = \min \underline{c}^T\underline{x} + \underline{u}^T(\underline{d} - D\underline{x}) \quad \text{st } \underline{x}\in X\subseteq\mathbb{R}^n \qquad (3) \text{ Lagrangian sub-problem}$$

$$w^* = \max w(\underline{u}) \quad \text{st } \underline{u}\geq \underline{0} \qquad (4) \text{ Lagrangian dual}$$

So we get that (3) is a relaxation of (2) for any multiplier vector $\underline{u}\geq 0$. And (4) is to ask to find the most tight lower bound (ie the closest to z^*). Otherwise being a relaxation we may end up with $w(\underline{u})$ a lot lower than (to) z^*.

Cor. (weak duality)

$\underline{x}$ feasible sol of (2), $\underline{u}$ feasible sol of (4) ⇒ $w(\underline{u}) \leq \underline{c}^T\underline{x}$



⇒ if $w(\tilde{\underline{u}}) = \underline{c}^T\tilde{\underline{x}}$, then $\tilde{\underline{x}}$ and $\tilde{\underline{u}}$ are both optimal, resp for (2) and (4)

⇒ if one problem is unbounded, the other is infeasible

If we Lagrangianize on equality constr, we get the (4) with free variables $\underline{u}$, ie $\underline{u}\gtrless 0$.

To Lagrangianize a problem:

- select a complicating constr we wish to remove
- see for how much times it occurs (like "$\forall i\in M$")
- introduce $\underline{u}$ of cardinality the above one (eg u_i, $i\in M$)
- remove the constr and update the obj function

Prop. If $\underline{u}\geq \underline{0}$ and

- $\underline{x}(\underline{u})$ is an opt sol of (3)
- $D\underline{x}(\underline{u}) \geq \underline{d}$ (ie satisfies removed constr)
- $[D\underline{x}(\underline{u})]_i = d_i$ for each $u_i > 0$ (constr met at equality for strictly positive u_i's)

⇒ $\underline{x}(\underline{u})$ is also opt for (2)

We also observe that the function $w(\underline{u})$ is concave (in this case of problems being min - min - max).

STRENGTH AND CHOICE OF LAG DUAL (4)

Thm (characterization in terms of an LP).
Consider the problems (2), (3) and (4) with $X = \{\underline{x}\in\mathbb{Z}^n : A\underline{x}\geq \underline{b}\}$
Then we have that

$$w^* = \min \underline{c}^T\underline{x} \quad \text{st } D\underline{x}\geq \underline{d},\ \underline{x}\in \mathrm{conv}(X)$$

original problem (2) (with min and st $\underline{x}$ structure)

we retrieve the complex constr

staying in the convex hull of the remaining constr ($A\underline{x}\geq\underline{b}$ and $\mathbb{Z}^n$)

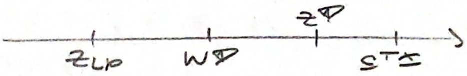
opt sol of the Lag dual

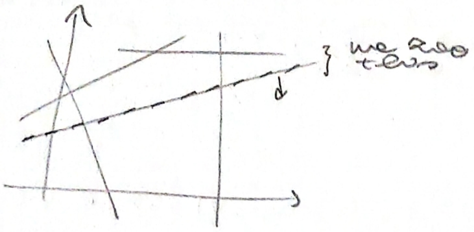
Cor. Since $\text{conv}(X) = \text{conv}(\{x \in \mathbb{Z}^n : Ax \geq b\}) \subseteq \{x \in \mathbb{R}^n : Ax \geq b\}$ we have

$z_{LP} \leq w^* \leq z^*$ — opt val of (2)

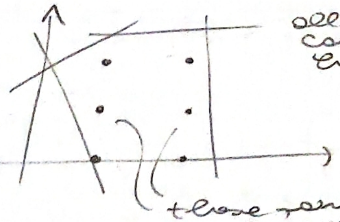
w^*: opt val of (a) → thus we have a better LB wrt LP relax

z_{LP}: opt val of the LP relaxation of (2)





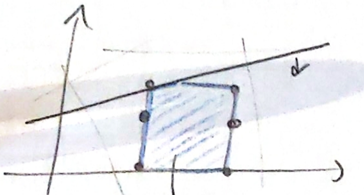
→ and we are left with



all the removed constr are from $Ax \geq b$

these points are the set $X = \{x \in \mathbb{Z}^n : Ax \geq b\}$

so we re-introduce it and we get ←



$\{x \in \text{conv}(X) : Dx \geq d\}$

the theorem says that w^* can be found by
- reintroducing $Dx \geq d$
- minimizing wrt c^Tx
- over conv(X) first and the constr $Dx \geq d$

Cor. Since the theorem talks about optimizing on $\{x \in \text{conv}(X) : Dx \geq d\}$, then if conv(X) is equal to the LP relaxation of the overall problem (2), we could have the Lag relaxation as weak as the LP relaxation.

Which constraint(s) should we dualize? we have to consider:
- the strength of the bound w^*
- the difficulty of solving (3) and (a)

SOLVING LAG DUALS

We use a generalization of the gradient method (for C^1 functions) to convex piecewise C^1 ones.

In general, the subgradient method is:

for $\min f(x)$
$\quad$ st $x \in \mathbb{R}^n$

- start from initial x_0
- at k-th iteration consider $\gamma_k \in \partial f(x_k)$ and update
$$x_{k+1} = x_k - \alpha_k \gamma_k \quad (\alpha_k > 0)$$
- we dont optimize wrt α_k since for not diff. C^1 functions the direction $-\gamma_k$ is not nec a descent one

Under some assumptions (f convex, unbounded for $\|x\| \to +\infty$, the $\alpha_k \to 0$ but not too fast, ie $\sum \alpha_k = +\infty$), we have that the subgradient method
- terminates after a finite # of iterations and converges to x^*
- or admits a subsequence from $\{x_k\}$ that converges to x^*

The Lagrangian dual was: $\max w(u)$ st $u \geq 0$ with $w(u)$ concave and piecewise linear.

How do we find subgradients of $w(u)$? There is a simple result:

Prop: Consider $\underline{\tilde{\mu}} \geq \underline{0}$ and $X(\underline{\tilde{\mu}})$ is the set of optimal sols of (3) which is

$$W(\underline{\tilde{\mu}}) = \min \underline{c}^T \underline{x} + \underline{\tilde{\mu}}^T(\underline{d} - D\underline{x})$$
$$\text{st } \underline{x} \in X \subseteq \mathbb{R}^n$$

$\Rightarrow$ - for each $\underline{x}(\underline{\tilde{\mu}}) \in X(\underline{\tilde{\mu}})$, the vector $(\underline{d} - D\underline{x}(\underline{\tilde{\mu}})) \in \partial W(\underline{\tilde{\mu}})$ ie the constr violation vector is a subgradient of $W(\underline{\tilde{\mu}})$

- the subgradients derived in this way span all the possible subgradients of $W(\underline{\mu})$ at $\underline{\tilde{\mu}}$.

So the procedure is:

- select initial $\underline{\mu}_0$ and set $k=0$.
- solve (3) to get $W(\underline{\mu}_k)$. If $\underline{x}(\underline{\mu}_k)$ opt sol is found, then $(\underline{d} - D\underline{x}(\underline{\mu}_k))$ is a subgradient of $W(\underline{\mu})$ at $\underline{\mu}_k$
- update Lagrange multipliers as

$$\underline{\mu}_{k+1} = \max(\underline{0}, \underline{\mu}_k + \alpha_k(\underline{d} - D\underline{x}(\underline{\mu}_k)))$$
(component-wise

- set $k=k+1$ and repeat.

UNCONSTRAINED NONLINEAR OPT

4.1-2 OPTIMALITY CONDITIONS

We have a generic problem:

$$\min_{\underline{x}\in S} f(\underline{x}) \quad \text{with} \quad S\subseteq\mathbb{R}^n,\ f: S\to\mathbb{R},\ f\in C^1 \text{ or } C^2$$

Def. $\underline{d}\in\mathbb{R}^n$ is a feasible direction at $\bar{\underline{x}}\in S$ $\Leftrightarrow$ $\exists\bar{\alpha}>0: \bar{\underline{x}}+\alpha\underline{d}\in S \ \forall\alpha\in[0,\bar{\alpha}]$

(1st order NEC local) $f\in C^1(S)$, $\bar{\underline{x}}$ loc min ($\in S$) $\Rightarrow$ $\forall\underline{d}\in\mathbb{R}^n$ feasible direction at $\bar{\underline{x}}$, $\nabla f(\bar{\underline{x}})\cdot\underline{d}\geq 0$

$\Rightarrow$ all feasible directions have to be ascent (worse) directions

(2nd order NEC local) $f\in C^2(S)$, $\bar{\underline{x}}$ loc min ($\in S$) $\Rightarrow$ (1) $\nabla f(\bar{\underline{x}})\cdot\underline{d}\geq 0$ $\forall$ feas. dir $\underline{d}$ at $\bar{\underline{x}}$
(2) if $\nabla f(\bar{\underline{x}})\cdot\underline{d}=0$ then (for that $\underline{d}$) $\underline{d}^T\nabla^2 f(\bar{\underline{x}})\,\underline{d}\geq 0$

$\downarrow$

Cor. $f\in C^2(S)$, $\bar{\underline{x}}$ loc min $\in \operatorname{int}(S)$ $\Leftrightarrow$ (1) $\nabla f(\bar{\underline{x}})=\underline{0}$ (classical stat condition)
(2) $\nabla^2 f(\bar{\underline{x}})$ is positive semi-definite

$\downarrow$

(SUF loc) $f\in C^2(S)$, $\bar{\underline{x}}\in\operatorname{int}(S)$ is st $\nabla f(\bar{\underline{x}})=\underline{0}$ and $\nabla^2 f(\bar{\underline{x}})$ is pos definite $\Rightarrow$ $\bar{\underline{x}}$ is a strict loc min, ie $f(\bar{\underline{x}})<f(\underline{x})$ $\forall\underline{x}\in N_\varepsilon(\bar{\underline{x}})\cap S$

CONVEX PROBLEMS

Now the problem definition is:

$$\min_{\underline{x}\in C} f(\underline{x}) \quad \text{with} \quad C\subseteq\mathbb{R}^n \text{ convex set},\ f: C\to\mathbb{R} \text{ convex function}$$

(NEC & SUF) f convex, C convex, $f\in C^1(C)$

$\underline{x}^*$ is a (global) min of f on C $\Leftrightarrow$ $\nabla f(\underline{x}^*)\cdot(\underline{x}-\underline{x}^*)\geq 0$ $\forall\underline{x}\in C$

The proof relies on the previous general result we stated above about 1st order NEC, and the characterization of convex function for SUF.

There is also this

Property. f convex, C convex bounded and closed.

f has finite maximum over C $\Rightarrow$ $\exists$ an optimal extreme point of C

4.3-4 ITERATIVE METHODS & CONVERGENCE

Now consider a general nonlinear optimization problem:

$$\boxed{\begin{array}{l}\min f(\underline{x}) \\ \text{st } g_i(\underline{x})\leq 0 \quad \forall i=1,\dots,m \\ \underline{x}\in S\subseteq\mathbb{R}^n\end{array}}\ , \quad f \text{ and } g_i\in C^1 \text{ at least}$$

$X=\{\underline{x}\in S: g_i(\underline{x})\leq 0\ \forall i\}$ is the feasible region

Most NO methods are iterative, ie:

- start from a certain $\underline{x}_0\in X$
- generate a sequence $\{\underline{x}_k\}_{k\geq 0}$ that converges (in some sense) to a point $\in\Omega$ the set of desired sets (eg points satisfying the above NEC optimality conditions)

We care about iterative methods being

(1) Robust, where we we want global convergence.

Def. a method is globally convergent if we reach Ω $\forall x_0 \in X$.
Otherwise, eg if only for $x_0 \in N_\varepsilon(x^*)$ with $x^* \in \Omega$, is locally conv.

(2) Efficient, ie fast numerically.

Def. We say that

$\{x_k\}$ converges to x^* with order $p \geq 1$ $\iff$ $\exists r>0$ (the rate) and $k_0 \in \mathbb{N}$ st
$\|x_{k+1}-x^*\| \leq r \cdot \|x_k - x^*\|^p$ $\forall k \geq k_0$

$\Rightarrow$ we look for the largest p and smallest r

About the rate r we know that convergence (if $p=1$) is
- linear if $r<1$
- sublinear if $r \geq 1$
- superlinear if r is actually r_k with $r_k \to 0$ as $k \to +\infty$.

LINE SEARCH METHODS

Now consider an unconstrained opt problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{with} \quad f: \mathbb{R}^n \to \mathbb{R},\ C^1 \text{ or } C^2, \text{ and bounded by below}$$

The general scheme of these methods is:

```
- select $x_0$ and $\varepsilon>0$, set $k=0$.
while (termination criterion is not satisfied)
    - choose search direction $d_k \in \mathbb{R}^n$
    - determine step length $\alpha_k$ along $d_k$ such to
          $\min_{\alpha \geq 0} \varphi(\alpha) = f(x_k + \alpha d_k)$
    - update $x_{k+1} = x_k + \alpha_k d_k$ and $k = k+1$
end
```

where the termination criterion is eg when $\|\nabla f(x_k)\| < \varepsilon$ or $|f(x_k) - f(x_{k+1})| < \varepsilon$ or $\|x_{k+1} - x_k\| < \varepsilon$.

About the search direction d_k we would like it to be a descent direction, and so a general structure is

$$\boxed{d_k = -D_k \nabla f(x_k)}$$ with D_k pos def (a matrix $n \times n$)

About the step length is sufficient to solve approximately that line search problem for α_k. Morally, it shouldn't be too small neither too large. So we get the following

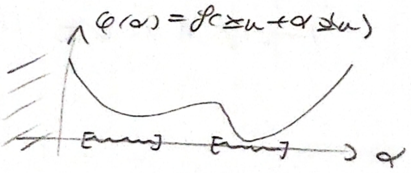
$$\boxed{\begin{aligned} \varphi(\alpha) &\leq \varphi(0) + c_1 \alpha \varphi'(0) & (1) \\ \varphi'(\alpha) &\geq c_2 \varphi'(0) & (2) \end{aligned}}$$

Wolfe conditions
(strong if (2) has $|\cdot|$ on the LHS)

with $0 < c_1 < c_2 < 1$

$$\iff \begin{aligned} & f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f(x_k) \cdot d_k \\ & d_k \cdot \nabla f(x_k + \alpha_k d_k) \geq c_2\, d_k \cdot \nabla f(x_k) \end{aligned}$$

The idea is that
(1) ensures that α_k makes f decrease sufficiently
(2) ensures that α_k makes the slope $\nabla f(x_k)$ to decrease sufficiently



Some theory guarantee that we can indeed find (under reasonable assumptions) some α_k satisfying these conditions, and also in finitely many steps.

Finally, about global convergence, we have a result showing it can be guaranteed under suitable ass on α_k and d_k.
Consider

$$\cos(\theta_k) = \begin{pmatrix} \text{angle between} \\ d_k \text{ and } -\nabla f(x_k) \end{pmatrix} = \frac{(-\nabla f(x_k)) \cdot d_k}{\|\nabla f(x_k)\| \cdot \|d_k\|}$$

Notably the global convergence holds when $\underline{d}_k$ is not too far away from the steepest descent direction $-\nabla f(\underline{x}_k)$. The fact $\cos(\theta_k) \geq \delta > 0$ i.e. $\theta_k \leq \pi/2$ i.e. $\underline{d}_k$ not $\perp$ to $-\nabla f(\underline{x}_k)$

θ_k
$-\nabla f(\underline{x}_k)$
$\underline{d}_k$
GOOD
$-\nabla f(\underline{x}_k)$
θ_k
$\underline{d}_k$
BAD

Thm (Zoutendijk) Consider
- any line search method with descent $\underline{d}_k$ and α_k satisfying Wolfe conditions
- f bounded below in $\mathbb{R}^n$, C^1 on N open set containing $L_0 = \{\underline{x} \in \mathbb{R}^n : f(\underline{x}) \leq f(\underline{x}_0)\}$ and $\nabla f(\underline{x})$ is Lipschitz continuous on N

Then

$$\Rightarrow \boxed{\sum_{k=0}^{+\infty} \cos^2(\theta_k) \cdot \|\nabla f(\underline{x}_k)\|^2 < +\infty}$$

$\Rightarrow$ this $\to 0$ if $\cos^2(\theta_k) \geq \delta > 0$ (as it is since this term has to go to zero)

and this can easily happen with more requirements on the matrix D_k of the $\underline{d}_k$ update

4.5 GRADIENT METHOD

Problem: given $f: \mathbb{R}^n \to \mathbb{R}$, C^1, look for a stationary point.
Gradient method: with exact 1D search we

```
choose $\underline{x}_0$, set $k=0$
while (! criterion met)
    set direction $\underline{d}_k = -\nabla f(\underline{x}_k)$
    find $\alpha_k$: $\min_{\alpha \geq 0} \varphi(\alpha) = f(\underline{x}_k + \alpha \underline{d}_k)$
    update $\underline{x}_{k+1} = \underline{x}_k + \alpha_k \underline{d}_k$, $k = k+1$
end
```

Property: if 1D search is exact, successive directions are orthogonal. This makes the method a bit slow.

What to monitor as criterion? the difference on the x-values or on the images is mostly equivalent:

Prop. If $H(\underline{x}^*)$ is pos def (strong min),
$\underline{x}_k$ converges (super) linearly to $\underline{x}^*$ wrt monotone $|f(\underline{x}_k) - f(\underline{x}^*)|$ $\Leftrightarrow$ it converges in the same way wrt $\|\underline{x}_k - \underline{x}^*\|$

QUADRATIC STRICTLY CONVEX FUNCTIONS

$$f(\underline{x}) = \frac{1}{2}\underline{x}^T Q \underline{x} - \underline{b}^T \underline{x} \quad \text{with } Q \text{ symm pos def}$$
$$(\nabla f(\underline{x}) = Q\underline{x} - \underline{b})$$

Here solving the 1D exact search problem for α_k is easy and we get

$$\alpha_k = \frac{\nabla f(\underline{x}_k)^T \nabla f(\underline{x}_k)}{\nabla f(\underline{x}_k)^T Q \nabla f(\underline{x}_k)} = \frac{\underline{d}_k^T \underline{d}_k}{\underline{d}_k^T Q \underline{d}_k}$$

Thm. For these functions, with exact 1D search, then $\forall \underline{x}_0$ we have that $\underline{x}_k \to \underline{x}_0$ (is globally convergent) with linear speed and a rate dependent on condition number of Q:

$$\|\underline{x}_{k+1} - \underline{x}^*\|_Q^2 \leq \left(\frac{\lambda(n) - \lambda(1)}{\lambda(n) + \lambda(1)}\right)^2 \|\underline{x}_k - \underline{x}^*\|_Q^2 \qquad (\|\underline{z}\|_Q^2 = \underline{z}^T Q \underline{z})$$

method is:
+ computationally light
+ globally convergent
− slow convergence

ARBITRARY FUNCTIONS

Thm. If $f \in C^2$, we use exact LS search, and we converge to x^* at $H(x^*)$ is pos def (here $H(x^*)$ makes the role of Q before) then we have a similar result but reversing the $f(\cdot)$ terms:

$$f(x_{k+1}) - f(x^*) \leq \left(\frac{\lambda(M) - \lambda(m)}{\lambda(M) + \lambda(m)} \right)^2 \cdot \left(f(x_k) - f(x^*) \right)$$

9.6 NEWTON METHOD

Problem: as before but now f is C^2.

(pure) newton method: let $H(x) = \nabla^2 f(x)$

- consider the quadratic approx of $f(x)$ at x_k:

$$g_k(x) := f(x_k) + \nabla f(x_k) \cdot (x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k)$$

- choose x_{k+1} as stationary point of $g_k(x)$:

$$\left.\begin{aligned} \nabla g_k(x) \overset{!}{=} 0 \iff& \nabla f(x_k) + H(x_k)(x - x_k) \overset{!}{=} 0 \\ \implies& x_{k+1} = x_k + \underset{\alpha_k = 1}{\uparrow} \underbrace{H(x_k)^{-1}(-\nabla f(x_k))}_{d_k} \end{aligned}\right\}$$ so this is exact search

This method is well defined if $H(x_k)$ is invertible, and that d_k will be a descent direction only if it is also pos def. There could be a α_k, but for pure newton we have fixed α_k.

method is:	− expensive computationally − only locally convergent + fast convergence + invariant wrt affine coordinates transformations

Thm. Suppose $f \in C^2$, x^* is st $\nabla f(x^*) = 0$ and $H(x^*)$ is pos def and Lipschitz continuous. Then

for x_0 sufficiently close to x^* $\implies$ (1) $x_k \to x^*$ quadratically
(2) $\|\nabla f(x_k)\| \to 0$ quadratically

MODIFICATIONS AND EXTENSIONS

(1) If $\alpha_k = 1$ does not satisfy Wolfe conditions we can set it [illegible] with inexact search.

(2) To guarantee to have d_k well defined and a descent direction we could set

not directly $d_k = H(x_k)^{-1}(-\nabla f(x_k))$

but rather $d_k = (\varepsilon_k I + H(x_k))^{-1}(-\nabla f(x_k))$

(3) using trust regions: we could limit the research of d_k and α_k just on a region B_k (adaptive) on which $g_k(x)$ approximates well $f(x)$.

CONJUGATE GRADIENT METHOD

Def. Let Q a symm n×n matrix. Two vectors $d_1, d_2 \in \mathbb{R}^n$ are Q-conjugate iff $\langle d_1, Q d_2 \rangle = d_1^T Q d_2 = 0$.

If Q is pos def then a set of vectors $d_0, \dots, d_k$ mutually Q-conj are also linear independent.

Consider the case of quadratic strictly convex functions, ie $g(x) = \frac{1}{2} x^T Q x - b^T x$.

Thm. Let $d_0, \dots, d_{n-1}$ be n Q-conjugate directions.

$\forall x_0$, the seq $\{x_k\}$ generated according to the exact line search scheme

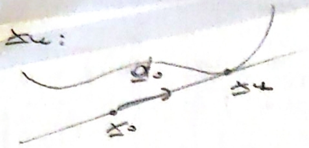
$$x_{k+1} = x_k + \alpha d_k, \quad \alpha_k = -\frac{\nabla g(x_k) \cdot d_k}{d_k \cdot (Q d_k)}$$

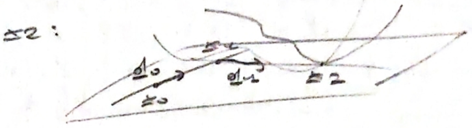
⇒ terminates to the unique global opt x^* in at most n iterations

$$x^* = x_0 + \sum_{k=0}^{n-1} \alpha_k d_k$$

Property: the optimization is "incremental", in the sense that $x^* = x_{k=n}$ is the global local optimum, but each of the x_k was the optimum on

- the half-previous line $\{x \in \mathbb{R}^n : x = x_{k-1} + \alpha d_{k-1}, \alpha \in \mathbb{R}\}$
- the full-previous subspace $\{x \in \mathbb{R}^n : x = x_0 + \sum_{i=0}^{k-1} \alpha_i d_i\} =: V_k$





⇒ $\nabla g(x_k)$ is ⊥ to V_k

How we derive Q-conjugate directions?

- initialization: x_0, $d_0 = -\nabla g(x_0)$, $k=0$
- iteration:

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

$$\alpha_k = \frac{-\nabla g(x_k) \cdot d_k}{d_k \cdot (Q d_k)}$$

(α_k comes from exact line search to minimize $g(x_k + \alpha_k d_k)$)

$$= \frac{d_k \cdot d_k}{d_k \cdot (Q d_k)}$$

$$d_{k+1} = \overbrace{-\nabla g(x_{k+1})}^{\text{this part resembles (is) the descent direction}} + \beta_k d_k$$

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

$$= (\text{proof}) \frac{\|\nabla g(x_{k+1})\|^2}{\|\nabla g(x_k)\|^2}$$

$$= \frac{\nabla g(x_{k+1})^T (\nabla g(x_{k+1}) - \nabla g(x_k))}{\|\nabla g(x_k)\|^2}$$

method is:

- + no computational heaviness
- - requires exact/accurate 1D search (the α_k) to not lose Q-conjugacy
- - not invariant wrt affine transformations

(in middle → faster convergence wrt gradient method
→ lower comp load wrt newton)

+ can be accelerated preconditioning

CONJUGATE DIRECTION METHODS

Now we are in the case of arbitrary functions.

- initialization: x_0, $d_0 = -\nabla f(x_0)$, $k=0$
- iteration

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

(α_k with inexact 1D search)

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

$$\beta_k^{FR} = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}$$

Fletcher-Reeves

$$\beta_k^{PR} = \frac{\nabla f(x_{k+1}) \cdot (\nabla f(x_{k+1}) - \nabla f(x_k))}{\|\nabla f(x_k)\|^2}$$

Polak-Ribiere

method vs

+ the real best choice for large n problems
+ small computationally cost
+ globally convergent if restart version (or even without)

Thm (convergence, CG): roughly, the more we iterate (we k↑), the better error estimates we get.

$$\underbrace{\|x_{k+1}-x^*\|_Q^2}_{e(k+1)} \le \left(\frac{\lambda_{n-k}-\lambda_1}{\lambda_{n-k}+\lambda_1}\right)^2 \underbrace{\|x_0-x^*\|_Q^2}_{e(0)}$$

$\lambda_1 \le \dots \le \lambda_n$ eigvals of Q

Thm (convergence, CD): convergence is super-linear when n iterations (under $f \in C^2$, FR method, exact search; but similar result hold for PR and inexact search).

QUASI NEWTON METHODS

Idea: rather than using $\nabla^2 f(x_k)$, we extract 2nd order information through variations of $\nabla f(x)$.

$x_{k+1} = x_k + \alpha_k d_k$ → Newton: $d_k = \boxed{(\nabla^2 f(x_k))^{-1}}(-\nabla f(x_k))$

→ Quasi N: $d_k = \boxed{H_k}(-\nabla f(x_k))$

⇒ we try to generate a sym pos def approx of the inverse of the hessian

method vs:

+ always well defined (and descent) directions
+ only involves 1st order derivatives
+ iteration cost $O(n^2)$ wrt $O(n^3)$ of Newton
− requires storing/handling matrices (disadvantage wrt CG or CD methods)

How do we characterize H_k? i.e. how do we derive the 2nd order information?

+ convergent wrt ... theory

+ less sensitive to inexact 1D search

+ can even ensure glo-bal conv

$$f(x_k+\delta) \approx f(x_k) + \delta^T \nabla f(x_k) + \frac{1}{2}\delta^T \nabla^2 f(x_k)\delta$$

$\frac{\partial}{\partial \delta}$

$$\nabla f(x_k+\delta) \approx \nabla f(x_k) + \nabla^2 f(x_k)\delta$$

↓ set $\delta_k := x_{k+1} - x_k$

$$\gamma_k := \nabla f(x_{k+1}) - \nabla f(x_k) \approx \nabla^2 f(x_k)(x_{k+1}-x_k)$$

$$\gamma_k \approx [\nabla^2 f(x_k)]\delta_k$$

$$[\nabla^2 f(x_k)]^{-1}\gamma_k \approx \delta_k$$

⇒ secant (*) condition: $H_{k+1}\gamma_k = \delta_k$

But in this way is not yet determined uniquely. So how do we update H_{k+1}?

- rank 1 update: $H_{k+1} = H_k + a_k u u^T$

(*), → $H_k\gamma_k + a_k u u^T \gamma_k = \delta_k$

multiplying by γ_k on both sides

$$\underbrace{a_k u (u^T\gamma_k)}_{1} = \delta_k - H_k\gamma_k$$

$\delta_k - H_k\gamma_k$ we can set those choices

$$a_k = \frac{1}{u^T\gamma_k}$$

Eventually this procedure will reach the true value (of $H_n = Q^{-1}$) but this rank 1 does not guarantee pos def so we move to

- rank 2 update: $H_{k+1} = H_k + a_k u u^T + b_k v v^T$

$$(*) \Rightarrow H_k \gamma_k + a_k u u^T \gamma_k + b_k v v^T \gamma_k = \delta_k$$

$$\underbrace{a_k}_{1}\underbrace{u}_{\delta_k}(\underbrace{u^T \gamma_k}_{d}) + \underbrace{b_k}_{-1}\underbrace{v}_{H_k\gamma_k}(\underbrace{v^T \gamma_k}_{d}) = \delta_k - H_k \gamma_k$$

$$a_k = \frac{1}{u^T \gamma_k} = \frac{1}{\delta_k \gamma_k} \qquad b_k = \frac{-1}{v^T \gamma_k} = \frac{-1}{(H_k \gamma_k)^T \gamma_k} = \frac{-1}{\gamma_k^T H_k \gamma_k}$$

We get in this way the DFP method.

Prop. If this holds:

$$\boxed{\text{Curvature (B) condition: } \delta_k \cdot \gamma_k > 0 \ \forall k} \Rightarrow$$ DFP method preserves pos def of H_k (ie if H_0 was def then the each ones will be)

But that condition is easy to prove, so we are good.

if (L) search satisfies Wolfe conditions $\Rightarrow$ (B) holds $\forall k \geq 0$

Towards BFGS method, the idea is now to construct approximations directly of the Hessian (rather than its inverse) but in a way that could be easily inverted.

So now for (*) we ask $\gamma_k = B_{k+1} \delta_k$ and again we can derive it incrementally through rank 2 updates.

Then the easy inversion comes from a result of linear algebra due to Sherman and Morrison.

BFGS has most of the same properties of DFP, but it's more robust in numerical terms.

Once we invert the BFGS matrix we can also combine it linearly with the DFP one to create the Broyden family of methods:

$$H_{k+1} = (1-\varphi) H_{k+1}^{DFP} + (\varphi) H_{k+1}^{BFGS}$$

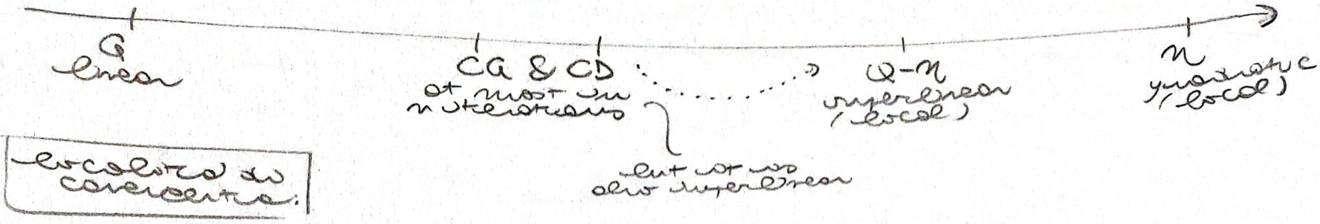
CONVERGENCE

Thm (Dennis and Moré) Roughly, we converge superlinearly iff (technical part: $f \in C^3$, B_k pos def, $\alpha_k = 1 \ \forall k$, x^* stationary point ie $\nabla f(x^*) = 0$ and local opt ie $D^2 f(x^*)$ pos def) the accuracy of B_k wrt the real Hessian $D^2 f(x^*)$ increases in the direction given by d_k:

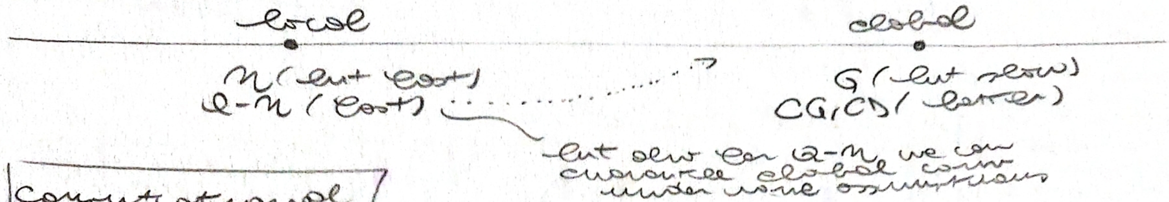
$$\lim_{k \to +\infty} \frac{\| (B_k - D^2 f(x^*)) d_k \|}{\| d_k \|} = 0$$

Thm. Under more assumptions, we can have global convergence even in the inexact (L) search case.

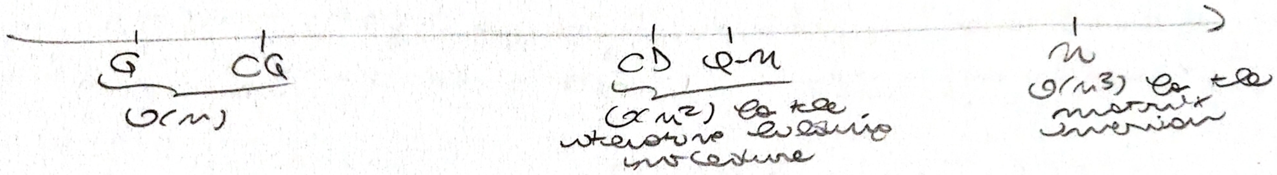
velocità di convergenza:



località di convergenza:


local
global
N (but fast)
Q-N (fast)
G (but slow)
CG, CD (better)
but also for Q-N we can guarantee global conv under some assumptions

computational cost:


G
CG
$O(n)$
CD
Q-N
$O(n^2)$ for the iterative updating procedure
N
$O(n^3)$ for the matrix inversion

CONSTRAINED NONLINEAR OPTIMIZATION

5.1 NEC OPT CONDITIONS

Consider the problem structure

$$\min f(x) \quad \text{st } g_i(x) \geq 0 \quad i \in I = \{1, \dots, m\}, \quad x \in \mathbb{R}^n$$

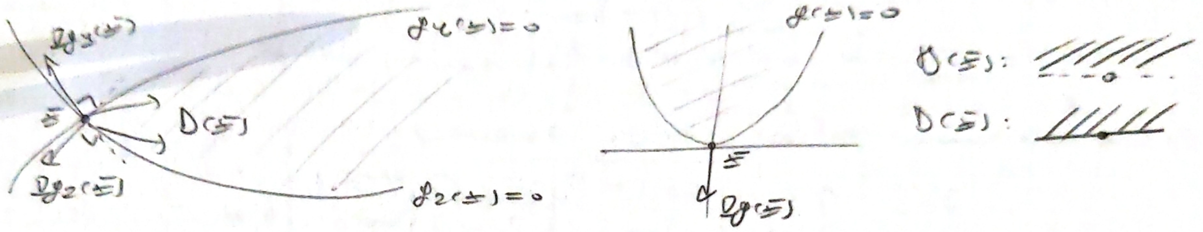
, $f, g_i \in C^1$, S the feasible region

Def. For $\bar{x} \in S$ let

$V(\bar{x}) = \{ \underline{d} \in \mathbb{R}^n : \exists \bar{\alpha} > 0 : x = \bar{x} + \alpha \underline{d} \in S \ \forall \alpha \in [0, \bar{\alpha}] \} =$
= cone of the feasible directions

$I(\bar{x}) = \{ i \in I : g_i(\bar{x}) = 0 \}$ = set of the active constr (ie the constr satisfied at eq)

$D(\bar{x}) = \{ \underline{d} \in \mathbb{R}^n : \nabla g_i(\bar{x}) \cdot \underline{d} \geq 0 \ \forall i \in I(\bar{x}) \} =$
= cone of the directions (not nec feasible!) constrained by the gradients of the active constraints



Property. $D(\bar{x}) \supseteq \overline{V(\bar{x})} \ \forall \bar{x} \in S$.

(1st order NEC) $f \in C^1(S)$, $\bar{x} \in S$ loc min $\Rightarrow \nabla f(\bar{x}) \cdot \underline{d} \geq 0 \quad \forall \underline{d} \in V(\bar{x})$
$\Rightarrow$ all feasible directions were ascent directions

not all vectors $\underline{d} \in D(\bar{x})$ are feasible directions, which are the ones of $V(\bar{x})$. However, this is difficult to characterize. So we ask the following:

Def (CQ, constraint qualification, Zangwill).

CQ condition holds at $\bar{x} \in S$ $\iff$ $\boxed{\overline{V(\bar{x})} = D(\bar{x})}$

if CQ ass does not hold, then KKT are not nec

Thm (KKT, nec opt conditions) Suppose $f \in C^1$, $g_i \in C^1$, CQ ass holds at $\bar{x} \in S$ (a feasible point).

$\bar{x}$ loc min of f on S $\Rightarrow$ $\exists$ KKT multipliers $u_1, \dots, u_m \geq 0$ st
$$\nabla f(\bar{x}) = \sum_{i \in I(\bar{x})} (-u_i) \nabla g_i(\bar{x})$$
active constr

$$\iff \begin{cases} \nabla f(\bar{x}) = \sum_{i=1}^m (-u_i) \nabla g_i(\bar{x}) \\ u_i g_i(\bar{x}) = 0 \quad \forall i = 1, \dots, m \end{cases}$$
all constr

So we will impose those conditions (KKT & slack) plus the ones of $\bar{x}$ satisfying the constr to have feasible candidate extremum points.

How can we ensure to have CQ ass satisfied?

Prop (suf conditions for CQ).

(1) g_i are linear $\forall i$, OR g_i convex $\forall i$ and $\exists$ an interior feasible point ($\hat{x} : g_i(\hat{x}) > 0 \ \forall i$) $\Rightarrow$ CQ holds $\forall \bar{x} \in S$
Karlin
Slater

(2) $\nabla g_i(\bar{x})$ are linearl. $\perp$ $\forall i \in I(\bar{x})$ (active constr!) $\Rightarrow$ CQ ass holds at $\bar{x} \in S$

Now the next result shows how the KKT conditions become suf for convex problems.

Theo (KKT nec and suf, convex problems). Suppose $f \in C^1$ convex, $g_i \in C^1$ are convex, and we have Slater cond (for CQ).

$x^* \in S$ is a glob min $\iff$ $\exists u_1, \dots, u_m \geq 0$ st
$$\begin{cases} \nabla f(x^*) = \sum_{i=1}^{m} (-u_i) \nabla g_i(x^*) \\ u_i g_i(x^*) = 0 \quad \forall i = 1, \dots, m \end{cases}$$

GENERAL CASE WITH EQUALITY CONSTR

Morally the equality constr will be always active constr, and then KKT multiplier will be free $\in \mathbb{R}$ (not in $\mathbb{R}^+$).

$$\begin{aligned} &\min f(x) \\ \text{st } & g_i(x) \leq 0 \quad i \in I \\ & h_\ell(x) = 0 \quad \ell \in L \\ & x \in X \subseteq \mathbb{R}^n \end{aligned} \quad , \quad \begin{aligned} & f, g_i, h_\ell \in C^1 \\ & S \text{ the feasible region} \end{aligned}$$

Due to equality, we usually get $D(\bar{x}) = \{0\}$, so we extend with a new definition.

Def.
$$T(\bar{x}) = \left\{ d \in \mathbb{R}^n : d = \lambda \cdot \lim_{k \to +\infty} \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|}, \lambda \geq 0, \{x_k\} \xrightarrow{k \to +\infty} \bar{x} \right\}$$
= closed cone of the tangents at $\bar{x}$

Def (CQ, constraint qualification, Abadie)

CQ condition holds at $\bar{x} \in S$ $\iff$ $\boxed{T(\bar{x}) = D(\bar{x}) \cap H(\bar{x})}$

this was the role of $D(\bar{x})$

$D = \{d : \nabla g_i(x) \cdot d \leq 0 \ \forall i \in I(\bar{x})\}$ active ineq constr

$H = \{d : \nabla h_\ell(x) \cdot d \leq 0 \ \forall \ell \in L\}$ "active eq constr" (are all of them)

Theo (KKT, nec opt conditions, general case) Suppose $f \in C^1$, $g_i \in C^1 \ \forall i$, $h_\ell \in C^1 \ \forall \ell$, and CQ cond holds at $\bar{x} \in S$.

$\bar{x}$ loc min of f on S $\implies$
$$\boxed{\begin{aligned} &\exists \boxed{u_i \geq 0} \ \forall i \in I(\bar{x}), \ \boxed{v_\ell \in \mathbb{R}} \ \forall \ell, \text{ st} \\ &\nabla f(\bar{x}) = \sum_{i \in I(\bar{x})} (-u_i) \nabla g_i(\bar{x}) + \sum_{\ell \in L} (-v_\ell) \nabla h_\ell(\bar{x}) \end{aligned}}$$

Prop (suf cond for CQ).

(1) g_i convex, h_ℓ linear, $\exists$ interior feasible point x for all g_i and h_ℓ $\implies$ CQ cond holds $\forall \bar{x} \in S$

(2) $\nabla g_i(\bar{x}) \ \forall i \in I(\bar{x})$ now are lin $\perp$ to $\nabla h_\ell(\bar{x}) \ \forall \ell \in L$ $\implies$ CQ cond holds at $\bar{x} \in S$

5.2 SUF OPT CONDITIONS

$$\begin{aligned} &\min f(x) \\ \text{st } & g_i(x) \leq 0 \quad i \in I = \{1, \dots, m\} \\ & x \in X \subseteq \mathbb{R}^n \text{ even discrete} \end{aligned} \qquad (P)$$

Def The Lagrange function associated to problem (P) is
$$L(x, u) = f(x) + \underbrace{\sum_{i \in I} u_i g_i(x)}_{\text{some penalization like we saw in DOPT}} \quad \begin{pmatrix} x \in X \\ u \geq 0 \end{pmatrix}$$

A point $(\bar{x}, \bar{u})$ with $\bar{x} \in X$, $\bar{u} \geq 0$ (not feasible for L) is a saddle point if
$$\bar{x} = \arg\min_{x \in X} L(\cdot, \bar{u}) \quad \& \quad \bar{u} = \arg\max_{u \geq 0} L(\bar{x}, \cdot)$$

Prop (characterization)

$(\bar{x}, \bar{u})$ with $\bar{x} \in X$, $\bar{u} \geq 0$ is saddle pt of L $\iff$
(1) $L(\bar{x}, \bar{u}) = \min_{x \in X} L(x, \bar{u})$ ~ $\bar{x}$ min
(2) $g_i(\bar{x}) \leq 0 \ \forall i \in I$ ~ $\bar{x}$ is feasible
(3) $\bar{u}_i g_i(\bar{x}) = 0 \ \forall i \in I$ ~ compl cond (slack)

Thm (suff opt conditions).

$(\bar{x}, \bar{u})$ is a saddle point of $L(x,u)$ ⇒ $\bar{x}$ is a glob min of problem (P)

It becomes also nec for convex problems, where it also emerges a connection to the KKT multipliers.

in the sense that for non convex problems a saddle point may not exist

Thm. Suppose f is convex, g_i are convex (conv with ∃ an interior feasible point). Then, with also X convex,

$$\boxed{\text{(P) has an optimal sol } x^* \iff \exists\, \bar{u} \ge 0 \text{ st } (\bar{x}, \bar{u}) \text{ is a saddle point for } L(x,u)}$$

(the previous thm)

For convex problems, the KKT ensured that a point satisfying them, say x^*, was the optimal sol. Turns out that

$$\binom{\text{KKT}}{\text{mult}} = \left(\text{Lagrange mult at the saddle pt}\right)$$

$$\hookrightarrow L(x,u) = f(x) + \sum_{i \in I} \overset{\ge 0}{u_i}\, g_i(x)$$

$$\nabla f = \sum_{i \in I} (-\underset{\ge 0}{u_i}) \nabla g$$

$$\boxed{\begin{array}{l}\nabla_x L = 0 \\ \iff \text{KKT}\end{array}} \Rightarrow \nabla f(\bar{x}) = \sum_{i \in I} (-u_i) \nabla g_i(\bar{x})$$

5.6 LAGRANGIAN DUALITY

What we also saw in DOPT: the idea is that for a min problem (P) we can solve it by solving the dual problem, looking for a saddle point of the Lagrange function.

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{st } g_i(x) \le 0 \quad i \in I \\ x \in X \subseteq \mathbb{R}^n \end{array} \qquad (D) \quad \max_{u \ge 0} \Big(\underbrace{w(u)}_{\min_{x \in X} L(x,u)} \Big)$$

⇒ we look for a saddle point

$$\begin{array}{l}\forall x \text{ feasible for (P)} \\ \forall u \text{ feasible for (D)}\end{array} \Rightarrow \Big|_{\text{weak duality}} \quad w(u) \le f(x)$$

under that, by strong duality:

$$w(\bar{u}) = f(\bar{x}) \iff \begin{array}{l}\bar{x} \text{ is opt for (P)} \\ \bar{u} \text{ is opt for (D)}\end{array}$$

$$\iff (\bar{x}, \bar{u}) \text{ is a saddle point}$$

So the method to solve exercise is to:
- derive the Lagrange function
- minimize it by deriving wrt x and setting it $=0$
- we find x in function of u, and we substitute those into the $L(x,u)$ to get $w(u)$, which we then maximize
- if no duality gap we should get $w(u^*) = f(x^*)$

When maximizing $w(u)$ in the last point remember that
(1) $w(u)$ is concave
(2) $g(x)$ is a subgradient of $w(u)$ if we solved the $\min_{x \in X} L(x,u)$ problem at optimality

In general (D) is easier than (P) even if no saddle point exists
- if saddle pt ∃: we can derive x^* optimal for (P) by the no-duality gap and requiring feasibility
- if saddle pt ∄: we can still find u^* by subgradient method, and the sequence of the $\{u_k\}$ will give a sequence of LBs $w(u_k)$ on the value $f(x^*)$, and we also create a sequence of $\{x_k\}$ by $\min_{x \in X} L(x, u_k)$.

Under convex assumptions there is no duality gap and there exists a saddle point.

5.5 2nd ORDER KKT CONDITIONS

$$\min f(\underline{x})$$
s.t. $g_i(\underline{x}) \leq 0 \quad i \in I \sim \underline{u}$
$h_e(\underline{x}) = 0 \quad e \in L \sim \underline{v}$, $f, g_i, h_e \in C^2$
$\underline{x} \in X \subseteq \mathbb{R}^n$ X open set

$$\Rightarrow L(\underline{x}, \underline{u}, \underline{v}) = f(\underline{x}) + \sum_{i \in I} \overset{\geq 0}{\underline{u_i}} g_i(\underline{x}) + \sum_{e \in L} \overset{\in \mathbb{R}}{\underline{v_e}} h_e(\underline{x})$$

(2nd ORD NEC KKT COND)

$\bar{x}$ loc min
$\nabla g_i \; \forall i \in I(\bar{x})$ are linear ⊥ to $\nabla h_e \; \forall e \in L$ $\Rightarrow$ ∃ some $(\bar{u}, \bar{v})$ that satisfy the KKT conditions

$\nabla_x L(\underline{x}, \underline{u}, \underline{v}) = 0$
$u_i g_i(\underline{x}) = 0$
$\underline{u} \geq 0, \; \underline{v} \in \mathbb{R}^{|L|}$ | KKT

$g_i(\underline{x}) \leq 0$
$h_e(\underline{x}) = 0$ | feas

Moreover, over $\underline{d} \in \mathbb{R}^n$ s.t.
$\nabla g_i(\bar{x}) \cdot \underline{d} \leq 0 \quad \forall i \in I(\bar{x})$
$\nabla h_e(\bar{x}) \cdot \underline{d} = 0 \quad \forall e \in L$
must satisfy $\underline{d} \cdot \nabla^2_{xx} L(\underline{x}, \bar{u}, \bar{v}) \underline{d} \geq 0$

(2nd ORD SUF KKT COND)

$\bar{x}$ satisfies with $(\bar{u}, \bar{v})$ the previous KKT cond
if $\underline{d} \cdot \nabla^2_{xx} L(\bar{x}, \bar{u}, \bar{v}) \underline{d} > 0$
$\forall \underline{d} \in \mathbb{R}^n$: $\nabla g_i(\bar{x}) \cdot \underline{d} = 0 \quad i \in I^+ = \{i : u_i > 0\}$
$\nabla g_i(\bar{x}) \cdot \underline{d} \leq 0 \quad i \in I^0 = \{i : u_i = 0\}$
$\nabla h_e(\bar{x}) \cdot \underline{d} = 0 \quad e \in L$

$\Rightarrow$ $\bar{x}$ is a strict loc min of f over X

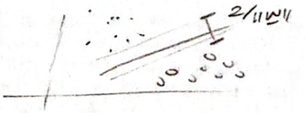
5.6 QUADRATIC PROGRAMMING

$$\min q(\underline{x}) = \tfrac{1}{2} \underline{x}^T Q \underline{x} + \underline{c}^T \underline{x}$$
s.t. $\underline{a}_i^T \underline{x} \leq b_i \quad i \in I$
$\underline{a}_i^T \underline{x} = b_i \quad i \in E$
$\underline{x} \in \mathbb{R}^n$

quadratic obj function, linear constraints, Q symm (wlog)

EX training (linear) SVM: find the separating hyperplane with the largest margin

$$\min_{\underline{x} = \binom{\underline{w}}{b} \in \binom{\mathbb{R}^n}{\mathbb{R}}} \tfrac{1}{2} \|\underline{w}\|^2$$
s.t. $y_i(\underline{w}^T \underline{x}_i - b) - 1 \boxed{\geq} 0$
$\forall$ points i } large number of constr $\Rightarrow$ we dualize



we need a ≤ 0 to dualize as we are used to

$$\max_{\underline{u} \geq 0} \Big(\min_{\underline{x} \in X} \underbrace{L(\underline{x} = \tbinom{\underline{w}}{b}, \underline{u})}_{[\frac{1}{2}\|\underline{w}\|^2] + \sum_i u_i [\ominus(y_i(\underline{w}^T \underline{x}_i - b) - 1)]} \Big)$$

QP WITH ONLY EQUALITY CONSTR

$\min \{ \tfrac{1}{2} \underline{x}^T Q \underline{x} + \underline{c}^T \underline{x} : A\underline{x} = \underline{b} \}$
so we just have (lin) sys to solve

$\underline{a}_i^T \underline{x} = \underline{b} \quad \forall i \in E$ — a_i rows are vectors these
$\Rightarrow A = \begin{pmatrix} - a_1 - \\ - a_m - \end{pmatrix}$

we write the KKT conditions:

$$\begin{cases} \underbrace{(Q\underline{x} + \underline{c})}_{\nabla f(\underline{x})} + \sum_{i=1}^{m} u_i \underbrace{(\underline{a}_i)}_{\nabla \text{constr}} = \underline{0} \\ A\underline{x} = \underline{b} \end{cases}$$

CQ satisfied from linear constr

compl slack are automatically satisfied

which is solved by this expanded linear system, derived from rewriting the conditions:

$$\begin{cases} Q\underline{x} + \sum \underline{a}_i \mu_i = -c \\ A\underline{x} = \underline{b} \end{cases} \Rightarrow \begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} \underline{x} \\ \underline{\mu} \end{pmatrix} = \begin{pmatrix} -\underline{c} \\ \underline{b} \end{pmatrix}$$

Null space method: we can find a matrix Z whose cols are ker(A) and then reformulate the variables as $\underline{x} = \underline{x}_0 + Z\underline{w}$ with
- $\underline{x}_0$ a feasible pt
- $\underline{w}$ an appropriate vector $\in \mathbb{R}^{n-m}$

In this way the problem gets converted to an unconstrained QP

$$\min_{w \in \mathbb{R}^{n-m}} \left[\frac{1}{2} w^T (Z^T Q Z) \underline{w} + (Q\underline{x}_0 + \underline{c})^T Z \underline{w} \right]$$

QP WITH INEQUALITY CONSTR

Knowing now the solvers of solutions for eq constr, the idea in this case is to:
- determine $I(\underline{x}^*)$, the active set around the opt set
- by solving a sequence of QPs with only eq constr

active set method: find an initial feasible $\underline{x}_0$ and choose $W_0 = E \cup \{i \in I : \underline{a}_i^T \underline{x}_0 = b_i\} = E \cup I(\underline{x}_0)$. Then at iteration k:
- determine $\underline{d}_k$ as follows

$$\min\{ \underbrace{q(\underline{x}_k + \underline{d})}_{\text{we minimize } q \text{ along direction } \underline{d}} : \underbrace{\underline{a}_i^T(\underline{x}_k + \underline{d}) = b_i,\ i \in W_k}_{\text{still satisfying all constr required by } W_k} \}$$

$$\Leftrightarrow \min\{ q(\underline{x}_k + \underline{d}) : \underline{a}_i^T \underline{d} = \underline{0},\ i \in W_k \}$$

- based on $\underline{d}_k$ we determine α_k to update $\underline{x}_{k+1}$ as $\underline{x}_k + \alpha_k \underline{d}_k$, and to update W_{k+1}.
  - if $\boxed{\underline{d}_k \neq \underline{0}}$ we find α_k as the largest step length st we satisfy all the feasibility constr (the ones outside W_k)

$$\Rightarrow \alpha_k = \min\left(1, \min_{\substack{i \notin W_k: \\ \underline{a}_i^T \underline{d}_k > 0}} \frac{b_i - \underline{a}_i^T \underline{x}_k}{\underline{a}_i^T \underline{x}_k}\right),$$

$$W_{k+1} = W_k \cup \{i'\}$$ with i' the index of the constr becoming active at $\underline{x}_{k+1}$

  - if $\boxed{\underline{d}_k = \underline{0}}$ then $\underline{x}_k$ is the min over the W_k constr. We need to check if it is the loc opt of the full problem: we compute the KKT multipliers

$$(Q\underline{x}_k + \underline{c}) + \sum_{i \in W_k} \mu_i^k \underline{a}_i = \underline{0} \sim \text{on the classical } \nabla f(\underline{x}_k) = \sum_{i \in W_k} (-\mu_i^k)\underline{a}_i$$

actually here the μ_i are the ones computed in the system when minimizing for $q(\underline{x}_k + \underline{d})$

$$\begin{cases} \text{if } \mu_i^k \geq 0\ \forall i \Rightarrow \underline{x}_k \text{ is the loc opt pt of QP} \\ \text{if } \exists i : \mu_i^k < 0 \Rightarrow \underline{x}_{k+1} = \underline{x}_k \text{ and } W_{k+1} = W_k \setminus \{i'\} \\ \qquad \text{where } i' \text{ the most neg index} \end{cases}$$

5.7 PENALTY METHOD & AUGMENTED LAG

Penalty method: the idea is to
- remove a constr
- penalize its/their violation in the obj function
- solve a sequence of unconstrained (QP) problems

$$(P) \quad \min_{\underline{x} \in \mathbb{R}^n} f(\underline{x}) \quad \text{st } c_i(\underline{x}) = 0 \quad i \in E$$

idea: let $\mu_k \to 0$ and find for each k the $\underline{x}_k$ minimizing that $Q(\cdot, \mu_k)$

(quadratic penalty-fnct problem)

$$\min_{\underline{x} \in \mathbb{R}^n} Q(\underline{x}, \mu) = f(\underline{x}) + \frac{1}{2\mu} \sum_{i \in E} c_i(\underline{x})^2$$

General scheme:

(0) select $\varepsilon>0$, $\gamma_0>>$, a sequence $\tau_k\to 0$, x_0^s, $k=0$

(1) compute x_k approximate minimizer of $Q(\cdot,\gamma_k)$ starting from x_k^s and terminating when $\|\nabla Q(x_k,\gamma_k)\|\le\tau_k$.

(2) if overall termination criterion is met (eg $|f(x_{k+1})-f(x_k)|<\varepsilon$)
then → return sol x_k
else → reduce $\gamma_{k+1}\in(0,\gamma_k)$, set $x_{k+1}^s=x_k$, $k=k+1$
and repeat from (1)

Thm (unrealistic)

x_k glob min of $Q(\cdot,\gamma_k)$ $\forall k$ (so $\tau_k=0$ $\forall k$), $\gamma_k\to 0$ ⇒ every limit point x^* of seq. x_k is a glob min of (P)

Thm (more realistic)

$\tau_k\to 0$, $\gamma_k\to 0$ ⇒ every limit point x^* at which $\nabla c_i(x^*)$ are lin ⊥ is a KKT point of (P)

Moreover, the sub-sequence defined by K with $x_k\to x^*$ is st

$$\lambda_i^* = \lim_{k\in K} -\frac{c_i(x_k)}{\gamma_k}$$

point λ^* that satisfies KKT cond with x^*. This is because

$$L(x,\lambda)=f(x)-\sum\lambda_i c_i(x) \Rightarrow \nabla_x L=\nabla f-\sum\lambda_i\nabla c_i \overset{!}{=} 0$$
$$Q(x,\gamma)=f(x)+\frac{1}{2\gamma}\sum c_i(x)^2 \Rightarrow \nabla_x Q=\nabla f+\frac{1}{\gamma}\sum c_i(x)\nabla c_i \overset{!}{=} 0$$

However, for $\gamma_k\to 0$ the quadratic penalty problem becomes ill conditioned. So the sol will be to move the quadratic penalization on the Lagrangian function.

Just to conclude, we can employ this method also with ineq constr by a term $1/2\gamma\sum[(c_i(x))^-]^2$, considering it only if $c_i(x)\ge 0$ ie. And we could also use other penalizations than the quadratic one.

Augmented Lagrangian method: reduce ill conditioning problems by
- moving the penalization on the Lagrangian
- introducing explicit estimates for the Lag multipliers

$$(P)\quad \begin{array}{l}\min f(x)\\ \text{st } c_i(x)=0 \quad i\in E\\ x\in\mathbb{R}^n\end{array}$$

$$(AP)\quad \min_{x\in\mathbb{R}^n} L_A(x,\lambda,\gamma)=\underbrace{f(x)-\sum_{i\in E}\lambda_i c_i(x)}_{\text{Lag function}}+\underbrace{\frac{1}{2\gamma}\sum_{i\in E}(c_i(x))^2}_{\text{quadratic penalty term}}$$

General scheme:

(0) initialize $\varepsilon>0$, $\gamma_0>0$, $\tau_k>0$: $\tau_k\to 0$, x^0, λ^0, $k=0$.

(1) determine approx minimizer x_k of $L_A(x,\lambda^k,\gamma_k)$ starting from x_k^s until $\|\nabla_x L_A\|\le\tau_k$. Note that

$$\nabla_x L_A(\dots)=\nabla f(x)-\sum_{i\in E}\underbrace{\left(\lambda_i-\frac{c_i(x)}{\gamma}\right)}_{\text{these will appear to be the optimal KKT }\lambda^*\text{ multiplier}}\nabla c_i(x)$$

(2) if overall termination criterion is met
then → stop
else → update $\lambda_i^{k+1}=\lambda_i^k-\frac{c_i(x_k)}{\gamma_k}$, reduce $\gamma_{k+1}\in(0,\gamma_k)$,
set $x_{k+1}^s=x_k$ and loop.

Thm: x^* loc min at which $\nabla c_i(x^*)$ are lin ⊥, 2nd order suff opt cond satisfied at x^*, λ^* and γ ⇒ the problem is well posed/defined (∃! set x_k at iter k, ecc) we have good conv speed.

5.8 BARRIER METHOD

$$(P) \quad \begin{aligned} \min\ & f(x) \\ \text{st}\ & c_i(x) \geq 0 \quad i \in I \\ & x \in \mathbb{R}^n \end{aligned}$$, let X be the feasible region

Def. Let $X^0 = \text{int}(X) \neq \emptyset$. A function $b: \mathbb{R}^n \to \mathbb{R}$ is a barrier function if it is continuous over X^0, $b(x) \xrightarrow[x \to \partial X]{} +\infty$, $b(x) = +\infty$ for $x \in \mathbb{R}^n \setminus X^0$.

So the idea of the method is:
- add to the obj function the barrier terms associated to the constr
- solve a sequence of unconstrained opt problems, like

(logarithmic barrier problem) $$\min_{x \in \mathbb{R}^n} P(x, \mu) = f(x) + \overset{>0}{\tilde{\mu}} \sum_{i \in I} (-\ln(c_i(x)))$$

General scheme: start from $x^0 \in X^0$ and for each k (and corresponding penalization coeff $\mu_k \to 0$) compute x_k approximate minimizer of $P(\cdot, \mu_k)$.

Obs: (1) starting from x_0 and penalizing moving out of it (or even just towards ∂X) we will always get interior pts

(2) relation to KKT conditions:

$$\text{KKT} \iff \nabla_x L \overset{!}{=} \underline{0} \iff \nabla f(x) - \sum_{i \in I} \lambda_i \nabla c_i(x) = \underline{0}$$

(−) since now (P) is defined with $\geq$ constraints, not $\leq$

min over x of $P(\cdot, \mu)$ $$\iff \nabla_x P \overset{!}{=} \underline{0} \iff \nabla f(x) - \sum_{i \in I} \left[\frac{\mu}{c_i(x)} \right] \nabla c_i(x) = \underline{0}$$

⇒ when $\mu \to 0$, a minimizer $x(\mu)$ of $P(\cdot, \mu)$ and the corresponding estimates $\lambda_i(\mu) = \mu / c_i(x(\mu))$ (that carry μ) tend to satisfy the KKT cond of (P)

Rem. Like above as before, under milder conditions all is good and well defined.

5.9 SEQUENTIAL QP

$$(P) \quad \begin{aligned} \min\ & f(x) \\ \text{st}\ & g_i(x) \leq 0 \quad i \in I \\ & h_l(x) = 0 \quad l \in E \\ & x \in \mathbb{R}^n \end{aligned}$$, f, g_i, h_l are C^2

performing a change of variables $x = x_k + \underline{d}$

Idea: extend Newton's method to nonlinearly constrained problems. So we derive the quadratic opt problem to find a direction d_k to improve the current x_k:

$$(Q-A_k) \quad \begin{aligned} \min\ & f(x_k) + \nabla f(x_k)^T \underline{d} + \tfrac{1}{2} \underline{d}^T \nabla^2 f(x_k) \underline{d} \\ \text{st}\ & Qg_i(x_k) \leq 0 \quad i \in I \\ & Qh_l(x_k) = 0 \quad l \in L \end{aligned} \Big\} \text{ quadratic approx of the constr}$$

This is difficult having quadratic constr, so the idea/alg is to
- consider linear approx of the constr
- solving with the Lagrangian, where an opt point $(\underline{d}^*, \mu^*, \lambda^*)$ of the Lag of (QA$_k$) is also an opt pt of the following QP

$$(QP-A_k) \quad \begin{aligned} \min\ & \cancel{f(x_k)} + \nabla f(x_k)^T \underline{d} + \tfrac{1}{2} \underline{d}^T \nabla^2_{xx} L(x_k, \mu^*, \lambda^*) \underline{d} \\ \text{st}\ & g_i(x_k) + \nabla g_i(x_k)^T \underline{d} \leq 0 \quad i \in I \\ & h_l(x_k) + \nabla h_l(x_k)^T \underline{d} = 0 \quad l \in L \end{aligned} \Big\} \text{ linear approx of the constr}$$

⇒ method tries to solve sequentially, thus:
- carry out an iteration of Newton's method on the obj function
- enforce feasibility with the linearization of the constr