PROJECT ON QUADRATIC OPTIMIZATION.

November 20, 2012

PROJECT PROPOSAL: QUADRATIC OPTIMIZATION

GROUP MEMBERS: Bidisha Roy,Ankani Chattoraj,Sebanti Chakrabarti

- **Title**: Concave quadratic optimization.
- **Description**: The general quadratic problem consists of a quadratic objective functions and a set of linear inequality constraints as shown below:

$$\begin{aligned} \text{Minimize} \quad & Q(x) = C^T x + \frac{1}{2} x^T D x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

  where, C is an n vector b is an n vector. A is an $n \times n$ matrix and D is an $n \times n$ matrix.

  There are many types of quadratic programming problems, the classifications are based on the nature of the matrix D. We will mainly work on **concave quadratic problems** where D is **negative semi definite**.

- **Motivation**: Concave quadratic problems are of great importance in various fields of real life. This type of problems often arise in problems involving economies of scale. Certain aspects of VLSI chip design can also be formulated as concave quadratic problems. The classical quadratic assignment problem can also be formulated using this.

  Concave quadratic problems are the simplest among all quadratic problems. This has some nice properties e.g all solutions of a concave quadratic problem lie at some vertex of the feasible region.That is why solving concave quadratic problems is of great importance.

- **Problem definition**: Problems of optimizing a concave quadratic function with respect to set of linear constraints and non-negativity restrictions with a additional restriction that some or all of the variables are required to be integers. Relevant numerical examples will be provided.
- **Plan**:

- In our project we will first introduce what a quadratic optimization problem is and will give a brief description of various types of quadratic optimization problems where concave quadratic problems will be highlighted mostly.
- Then we will look into concave quadratic problems specifically. The properties of this type of problems which make it simple to solve will be discussed briefly.
- The final step will be about various methods of solving a concave quadratic problem e.g. **Extreme point ranking method, Cutting plane method, Solving by using convex envelope of the concave function, Solving the problem by reducing it to bilinear programming, Reduction to seperable form**. We will give brief descriptions of the methods mentioned here with some small examples.

- **References**: Wikipedia and some lecture notes.

PROJECT REPORT ONE: QUADRATIC OPTIMIZATION

GROUP MEMBERS: Bidisha Roy,Ankani Chattoraj,Sebanti Chakrabarti

A General Quadratic optimization problem can be written as:

$$\min f(x) = \frac{1}{2}X^TQX + CX$$

$$AX \leq b$$

$$X \geq 0$$

• C is a n-dimensional row vector describing the coefficients of the linear terms in the objective function.
• Q is a $n \times n$ symmetric matrix describing the coefficients of the quadratic terms, assumed without loss of generality.If not,replace it by $\frac{Q+Q^T}{2}$ which does not change the value of f.
• X is the n dimensional column vector and the decision variable,which if not non-negative,can be converted to non negative by a linear transformation.
• A is $n \times n$ matrix and b is an m-dimensional column vector,defining the constraints.
We assume that a feasible solution exists and that the constraints region is bounded.
If Q is positive definite then $f(x)$ is strictly convex for all feasible points and the problem has a unique local minimum which is also the global minimum.
Classification of quadratic optimization :
Quadratic problems are classified based on the nature of the matrix Q -
(1) Bilinear problems: When the matrix Q is such that there exists two sub vectors of distinct variables Y and Z of X so that the problem is linear when one of these vectors is fixed,then it is called a bilinear problem.
(2) Concave quadratic problems: When the matrix Q is negative semi definite that is all its eigenvalues are non positive,then it is called a Concave

Quadratic problem.
(3) Indefinite quadratic problem: When the matrix Q has both positive and negative eigenvalues,then it is called an Indefinite Quadratic problem.

$*$ **Karush-Kuhn-Tucker condition:** This condition is necessary in general case of quadratic optimization and it is also sufficient when Q is positive definite. The $lagrangianfunction$ for the quadratic optimization programme is-

$$L(X, \mu) = CX + \tfrac{1}{2}X^TQX + \mu(AX - b)$$

where μ is an m-dimension row vector.
The KKT conditions for a local minimum are given as follows :

$$\frac{\partial L}{\partial x_j} \geq 0, j = 1, ..., n, \qquad C + X^TQ + \mu A \geq 0 \tag{1}$$

$$\frac{\partial L}{\partial x_i} \leq 0, i = 1, ..., n, \qquad AX - b \leq 0 \tag{2}$$

$$x_j \frac{\partial L}{\partial x_j} = 0, j = 1, ..., n, \quad X^T(C + X^TQ + \mu A) = 0 \tag{3}$$

$$\mu_i g_i(X) = 0, i = 1, ..., m \tag{4}$$

$$x_j \geq 0, j = 1, ..., n \qquad X \geq 0 \tag{5}$$

$$\mu_i \geq 0, i = 1, ...m \qquad \mu \geq 0 \tag{6}$$

To put (1)-(6) into a more manageable form we introduce non negative surplus variable $Y \in \mathbb{R}^n$ to the inequalities in (i) and non negative slack variable $v \in \mathbb{R}^n$ to the inequalities in (ii) to obtain the equations-

$C^T + QX + A^T\mu^T - Y = 0$ and $AX - b + v = 0$

So,the KKT conditions become:

$$QX + A^T\mu^T - Y = -C^T \tag{7}$$

$$AX + v = b \tag{8}$$

$$X \geq 0, \ y \geq 0, \ \mu \geq 0, \ v \geq 0$$
$$Y^TX = 0, \ \mu v = 0$$

The first two expressions are linear equations and the third is the non-negativity conditions and the fourth one comes from complementary slackness.

• METHODS OF SOLUTION OF CONCAVE PROBLEMS:

- Extreme point ranking method: The solution of the concave problems lies at a vertex of a polytope. So the solution can be obtained by complete enumeration of the extreme points.But this method is not very useful in case of large problems.

- Cutting plane methods : The basic approach of this methods involved starting from a specific vertex of the polytope.The edges of the polyhedron issuing from the vertex are used to define a cone that contain a feasible region.The use of cuts is to successively reduced the cone and generate an auxiliary subproblem which is solved in the sub cone,that give rise to a new vertex point.This is basically an iterative method.

- Convex envelops: concave problems can be developed from the use of the convex envelop of the concave function. This method based on the theorem –
  If $F(x)$ is the convex envelope of a concave function $f(x)$ taken over a convex domain S then any point that globally minimizes $f(x)$ over S also minimizes $F(x)$ over S.

- Reduction to bilinear programming: The concave quadratic program can also be solved by reduction to an associated bilinear programming problem.

- Reduction to separable form : It is possible to reduced the concave quadratic problem to a separable quadratic form. The method is given the matrix Q,compute the real eigenvalues of $-Q$ and the corresponding eigenvectors so that $Q = UDU^T$,$U = diag[u_1, u_2, ..., u_n]$,the matrix of eigen vectors and $U^TU + I$ and $D = diag[\alpha_1, ..., \alpha_n]$,the eigenvalues. These values are computed thena multiple-cost row linear programme with $2n$ rows are solved.Using the solution theQP can be reformulated in a separable form:

$$Min \sum_{n}^{i=1} q_i(x_i)$$
$$s.t. A^{'}X \leq b^{'}$$
$$X \geq 0$$

  Where $q_i(x_i) = c_i x_i - \frac{1}{2}\alpha_i x_i^2$ and $A^{'}$ and $b^{'}$ are produced from A and b through linear transformation using U.

PROJECT REPORT 2: SOLVING QUADRATIC OPTIMIZATION
GROUP MEMBERS: Bidisha Roy,Ankani Chattoraj,Sebanti Chakrabarti.

The General Quadratic Programming Problem

The general quadratic problem consists of a quadratic objective functions and a set of linear inequality constraints as given below:

$$\begin{aligned} max(f(x)) &= cx - \frac{1}{2}x^TQx \\ \text{Such that, } Ax &\leq b \\ x &\geq 0 \text{The objective function is} f(x)cx - \frac{1}{2}x^TQx \\ &= \sum_{j=1}^{n} c_j x_j - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_i x_j \end{aligned}$$

where, q_{ij} are elements of Q and Q is symmetric. If $i = j$ then $x_i x_j = x_j^2$. So $i\frac{1}{2}q_{ij}$ is the coefficient of x_i^2. If $i \neq j$ then $\frac{1}{2}(q_{ij}x_i x_j +)$, so $-q_{ij}$ is the coefficient for the product of x_i and x_j (since $q_{ji} = q_{ij}$).

Some necessary definitions

- HYPERPLANE : A hyperplane is a concept geometry. It is a general section of a plane into a different number of dimensions. A hyperplane of an n-dimensional space in a flat subset with dimension n-1. By nature it separates the space into two spaces.
- HALF SPACES: The points that are not incident to the hyperplane and portioned into two convex sets such that any subspace connecting a point in one set to a point in the other must intersect the hyperplane. Then such sets called half spaces.
- POLYTOPE: In elementary geometry, a polytope is a geometric object with flat sides which exists in any general number of dimensions .A polygon is a polytope of two dimensions and so on in higher dimensions.
- HYPERRECTANGALE: In geometry, a hyper rectangle is the generalisation of a rectangle for higher dimensions, formally defined as the Cartesian product of intervals.

Motivations of Cutting-Plane techniques

We prefer Cutting-plane methods as they

- do not require differentiability of the objective and constraint functions
- can exploit certain types of structure in large and complex problems
- do not require evaluation of the objective and all the constraint functions at each iteration

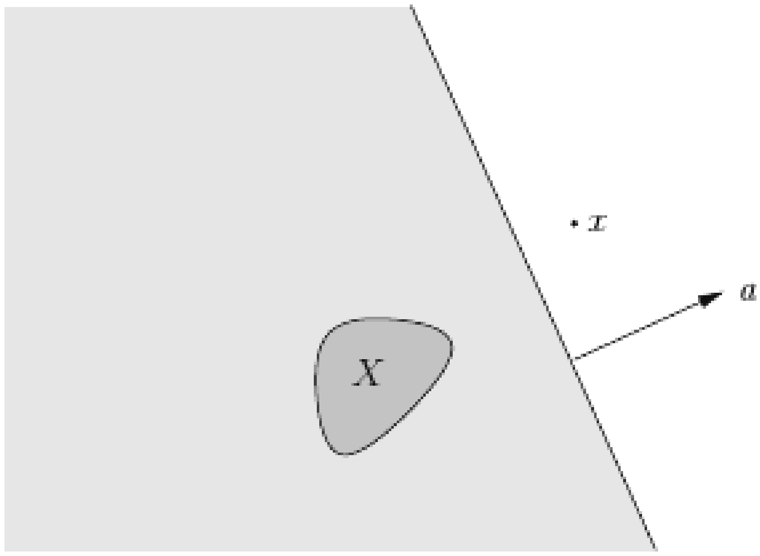
This can make cutting-plane methods useful for problems with a very large number of constraints.

Idea of Cutting-Plane technique

The goal of cutting-plane and localization methods is to find a point in a convex set $X \subset \mathbb{R}^n$, which we call the target set, (which might be empty set also). In an Optimization Problem, the target set X can be taken as the set of optimal points for the problem. To start, we do not have any direct description of the target set X.

In that case we proceed through an oracle (an approach to get an inequality). Suppose now we query the oracle at a point $x \in \mathbb{R}^n$, it returns the following information to us: it either tells us that $x \in X$ (in which case we are done), or it returns a separating hyperplane between x and X, i.e., $a \neq 0$ and b such that

$$a^T Z \leq b \text{ for } z \in X, a^T X \geq b$$



The inequality $a^T Z \leq b$ defines a cutting plane at the query point x, for the target set X, chosen shaded.

To find a point we need to search only the rectangle drawn above, the unshaded half space Z such that $a^T Z > b$ cannot contain any point in the target set.

Now, this hyperplane generated is called a cutting plane, or cut (as it eliminates the halfspace Z such that $a^T Z > b$ from our sketch as no such point can be in tangent set)

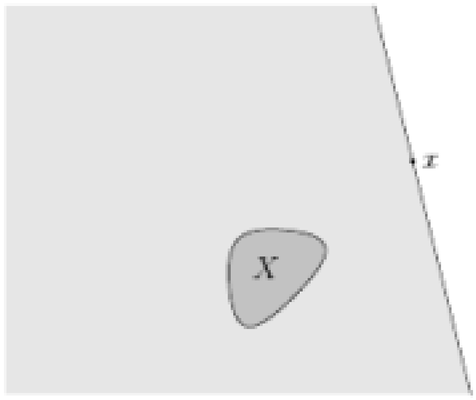
Now, there can be two cases-

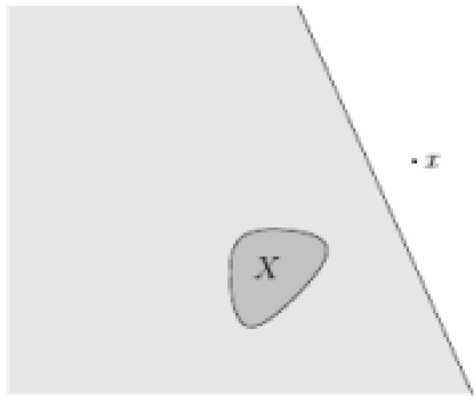
1. The cutting plane is generated on x (or the message that $x \in X$). When the cutting plane contains the query point x, we refer it as neutral cut or neutral cutting plane.

   Assumption: 2 norm of a is equal to 1 since dividing a and b by 2-norm of a defines the same cutting plane.

2. When $a^T \geq b$ which means that x lies in the interior of the halfspace that is being cut from consideration, the cutting plane is called deepcut.

   REMARK: A deepcut is better i.e. more informative than a neutral with the same normal vector excludes a larger set of points from consideration.





Tuy's Cutting-Plane

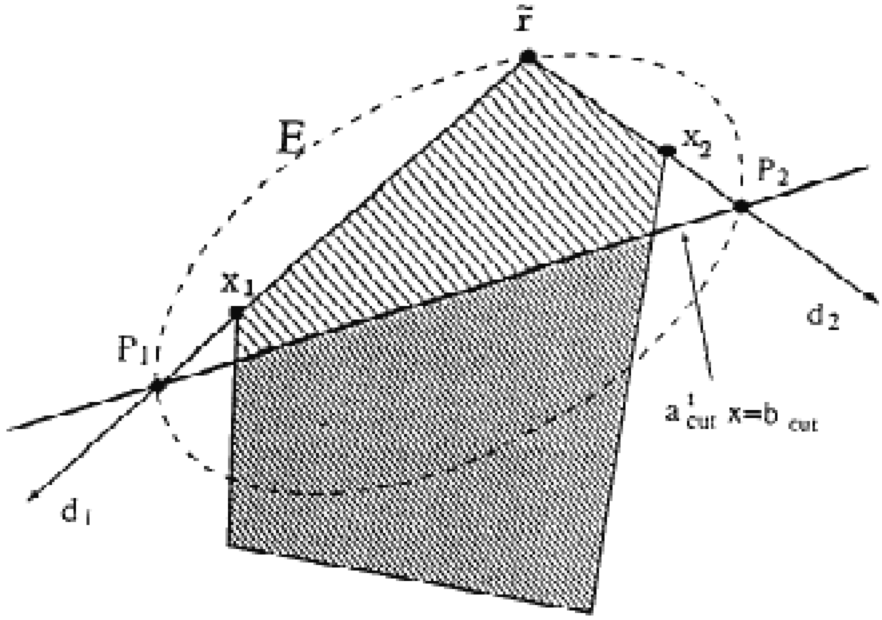
Pre requisite result:

1. 1) A global minimum of a concave function f over X exists among extreme points (vertices) of X. [**Journal of global optimization 13:225-240, 1998**]

2. A cutting plane is a linear constraint which eliminates a locally optimal solution and yet does not eliminate a globally optimal solution. [**Journal of global optimization 13:225-240, 1998**]

Method

Tuy is 1964 proposed a simple algorithm based on above ideas of cutting plane to solve a concave quadratic problem.



Let $x^{'}$ be a given locally optimal extreme point at which the value of the objective function is smaller than those at neighbouring vertices. Also, let E be an ellipsoid associated with the contour of $f(x^{''})$ where $x^{''}$ is the incumbent solution. Since $f(x)$ is a concave function, we have $f(x) \geq f(x^{''})$ for all x in E. Therefore, we can ignore the region X intersection E in the process of finding a global minimum of f over X. Along the edge d_i from $x^{'}$, we look for a point P_i at which the objective function value is equal to the value $f(x^{''})$. A linear constraint $a^t_{cut} \leq b_{cut}$ determined by $P_i, i = 1, 2, \ldots, n$ is the so-called Tuy's cutting plane.

REMARK:

1. Tuy's cut would be substantiality deeper when the objective function $f(.)$ is a low rank concave quadratic function. [**Journal of global optimization 13:225-240, 1998**]

2. We usually have a deeper cut when p is substantiality smaller than n. Where $f(.)$ in rank p concave quadratic programming problem. [**Journal of global optimization 13:225-240, 1998**]

Extension of this method

Rosen's cutting plane is an extension of this method. Tuy's cutting plane eliminates a portion of a feasible region in the neighbourhood of a locally optimal extreme point. Thus it may be called a neutral or boundary cut. Rosen came up in 1983 with Rosen cutting plane which occurs interior cut and eliminates a hyper rectangle whose centre is located at the global maximum of the objective function under the assumption that the feasible region is compact.

Cutting-Plane TUY/TABU search algorithm

TABU SEARCH: Tabu search is a local search algorithm that can be used for combinational optimization problems. Tabu search used a local neighbourhood search process to iteratively move from one potential solution x to an improved solution "x'' in the neighbourhood of x until some stopping criterion has been satisfied.

Now, we show the algorithm, for solving a concave quadratic programming problem.

$$\text{Minimize } c^t x + \frac{1}{2}x^t Qx$$
$$\text{Subject to } Ax \leq b$$

where, $c \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

We assume that the rank p of the non-positive definite symmetric matrix Q is small compared to $'n'$ and feasible region $X_0 = \{x \in \mathbb{R}^n \ni Ax \leq b\}$ is bounded.

We first solve a series of linear programming problem:

$$(LP)_K \text{such that} \qquad \text{Minimize } (c^k)^t x$$
$$\text{Subject to } Ax \leq b$$

where, c^k, $K = \{1, 2, \cdots, k\}$ are given set of vectors in $\mathbb{R}^n$. Let x^k be an optimal solution of $(LP)_k$ and let $V = \{x^1, x^2, \ldots, x^k\}$.

- We will choose a sequence of vectors c^k such that V is well scattered over X_0.
- Let $x_0 \in V$ be the starting point of cutting plane/Tabu search method, then the Tabu list consists of those points,
    1. which have been dominated by Tuy's cuts added during the course of computation,
    2. which are the set of extreme points of X_0 newly generated by Tuy's cuts which do not belong to V_0, the set of extreme points of X_0,
- These vertices satisfying condition (1) cannot be superior to the solution by the definition of Tuy's cut, and,
- These vertices satisfies (2) can be ignored in the search process became there is at least one optimal solution in V_0

Summary of the Topic

The advantage of such an approach is that at each iteration, the auxiliary problem has the same set of constraints and differs only in the objective function. Moreover the objective function at each step can be obtained from the previous iteration by simple linear transformation and column replacement.

NOTE: This method relies on all the vertices to be non degenerates. In case of degeneracy a suitable perturbation methods needs to be used.

However, Zwat(1973) proposed that the approach can be non-convergent due to cycling the need of an infinite sequence of cutting planes. The reason for cycling behaviour as well as nonconvergence of these approaches lies in the fact that although the approaches generate cones during the algorithm, they failed to explicitly incorporate these cones into the remaining steps. This is essential to avoid the re-emergence of the vertices that are already been considered. This difficulty with the approaches was recognized by Zwat(1974),who proposed a modified e-convergent approach where, at each iteration, constraints are added to ensure that the solution of the LP at each iteration is contained in the cone with the vertex at the current local minimum and a perturbed set of extreme rays coincident at the vertex.

Books

Nonlinear optimization - Andrzej and Ruszczy Ski

FINAL PROJECT REPORT: SOLVING QUADRATIC OPTIMIZATION

GROUP MEMBERS: Bidisha Roy,Ankani Chattoraj,Sebanti Chakrabarti.

Scalar Quadratic optimization without constraints:
minimize($\frac{1}{2}qx^2 + cx + c_o$) has a finite solution if and only if

- $q = 0$ and $c = 0$ this is the one condition.
- The another condition is $q > 0$

In the first case there is no unique optimum.
In the second case we can write-
$\frac{1}{2}\ qx^2 + cx + c_o \ = \ \frac{1}{2}q(x + \frac{c}{q})^2 + c_o - \frac{c^2}{2q}$ and the optimum is achieved for $x = \frac{-c}{q}$

Now another way for solving this problem is to factorize Q in the form

$$Q = LDL^T$$

Then,

$$\begin{aligned}
&\min \frac{1}{2}X^T QX + C^T X \\
=&\min \frac{1}{2}X^T LDL^T X + C^T (L^T)^{-1} L^T X \\
=&\min \sum_{k=1}^{n} \frac{1}{2} d_k y_k^2 + \tilde{C}_k y_k \\
=&\sum_{k=1}^{n} \min \frac{1}{2} d_k y_k^2 + \tilde{C}_k y_k
\end{aligned}$$

where $y = L^T X$ and $\tilde{C} = L^{-1}C$. Thus, we have reduced to the problem to n independent scalar quadratic Optimization Problems, which are now easy to solve.
Now $\hat{Q_1} = [\hat{q_1}, ..., \hat{q_n}]^T$ will be a minimizer when-

- $d_k \ \geq \ 0; k = 1, ...n$
  $\Leftrightarrow Q \ \ is \ \ a \ \ positive \ \ semi-definite.$

- and $d_k \hat{q_k} \ = \ -\tilde{c_k}$
  $\Leftrightarrow DL^T \hat{x} \ = \ -\mathrm{L}^{-1} c \quad \Longrightarrow LDL^T \hat{x} = c \quad i.e. \quad Q\hat{x} \ = \ -c$

Furthermore, $f(x) \ = \ \frac{1}{2} X^T Q X + C^T X + C_o$ **is unbounded if-**

- some $d_k \ < \ 0 \quad \Leftrightarrow$ Q is not positive semi-definite or
- some of the equations $d_k \hat{q_k} \ = \ -\tilde{c_k}$ do not have a solution i.e. $d_k \ = \ 0$ and $\tilde{c_k} \ \neq \ 0 \Leftrightarrow$ The equation $Q\hat{x} \ = \ -c$ do not have a solution.

Theorem: Let $f(x) \ = \ \frac{1}{2} X^T Q X + C^T X + C_o, \hat{x}$ will be a minimizer if,
(i) Q is a positive semi-definite.
(ii) $Q\hat{X} \ = \ -C$ i.e. $\hat{X} \ = \ -Q^{-1}C$
Comment: The condition that $QX \ = \ -C$ has a solution is equivalent to $c \ \in \ \Re(Q)$.
Quadratic optimization with Linear Constraints:

$$\begin{gathered} \text{minimize } \frac{1}{2} X^T Q X + C^T X + C_o \\ \text{such that } AX \ = \ b \end{gathered}$$

where $Q \ = \ Q^T \ \in \ \mathbb{R}^{n \ \times \ n}$,$A \ \in \ \mathbb{R}^{m \ \times \ n}$,$C \ \in \ \mathbb{R}^n$,$b \ \in \ \mathbb{R}^m$ and $C_o \ \in \ \mathbb{R}$.

Assumption:

- We assume that $b \ \in \ \Re(A)$ and $n \ > \ m$ i.e $N(A) \ \neq \ \{0\}$.
- Assume that $N(A) \ = \ span\{Z_1, ..., Z_k\}$ and define the null space matrix $Z \ = \ [Z_1, ..., Z_n]$.

If $A\bar{X} \ = \ b$,it holds that an arbitrary solution to the linear constraint has the form $X \ = \ \bar{X} + Zv$ for some $v \in \ \mathbb{R}^k$,This follows since,

$$A(\bar{X} + Zv) \ = \ A\bar{X} + AZv \ = \ b + 0 = b$$

with $X \ = \ \bar{X} + Zv$ inserted in the problem.We get

$$f(X) \ = \ \frac{1}{2} v^T Z^T Q Z v + (Z^T(Q\bar{X} + C))^T v + f(\bar{X})$$

We can now apply previous theorem to obtain the following, $\hat{X}$ is an optimal solution to

$$\begin{gathered} \text{minimize } \frac{1}{2} X^T Q X + C^T X + C_o \\ \text{such that } AX \ = \ b \end{gathered}$$

if- (i) $Z^T Q Z$ is positive semi definite.
(ii)$\hat{X} \ = \ \bar{X} + \hat{Z}v$ where $Z^T Q Z \hat{v} \ = \ -Z^T(Q\bar{X} + C)$

Comment: The second condition in the theorem is equivalent with the existence of a $\hat{u} \in \mathbb{R}^m$ such that

$$\begin{bmatrix} Q & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

The second constraint in the theorem can be written

$$Z^T \left(Q(\bar{x} + Z\hat{v}) + c \right) = 0$$

Since, $\mathcal{N}(A) = R(Z)$ this is equivalent with

$$Q(\bar{x} + Z\hat{v}) + c \in \mathcal{N}(A) = R(A^T)$$
$$\Longleftrightarrow Q\hat{x} + c = A^T\hat{u} \quad \text{for some} \quad \hat{u} \in \mathbb{R}^m$$

PROBLEM:

Let $f(x) \;=\; \frac{1}{2}x^TQx + C^Tx + C_o$, where $Q \;=\; \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $c \;=\; \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,$C_o = 0$. Consider the minimization of f over the feasible region $\xi \;=\; x \;\mid\; Ax = b$ where $A \;=\; [2 \;\; 1]$ and $b \;=\; 2$

Answer:

So the problem looks like:

$$\begin{bmatrix} min & -{x_1}^2 + \frac{1}{2}{x_2}^2 + x_1 + x_2 \\ & s.t.\;\; 2x_1 + x_2 \;=\; 2 \end{bmatrix}$$

The null space of A is spanned by Z and the reduced Hessian is given by,

$Z \;=\; \begin{bmatrix} -1 \\ 2 \end{bmatrix}$,$\bar{Q} = Z^TQZ \;=\; 2$

Since,the reduced Hessian is positive definite **so the problem is convex**.

Now the unique solution is then given by taking an arbitrary $\bar{x} \;\in\; \xi$ e.g. $\bar{x} \;=\; [1 \quad 0]^T$ and solving,

$\bar{Q}\hat{v} \;=\; -\bar{c} = -Z^T(Q\bar{x} + c) = -3$

$\therefore \hat{v} \;=\; -\frac{3}{2}$

$$\begin{aligned} \hat{x} &= \bar{x} + Z\hat{v} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}(-\frac{3}{2}) \\ &= \begin{bmatrix} -\frac{5}{2} \\ 3 \end{bmatrix} \end{aligned}$$