Optimization 1 - 098311

Winter 2021 - final exam

Asaf Gendler 301727715

February 10, 2021

Question 1:

a)

Let f be convex and continuously differentiable on a convex set C . Prove that x^* is an optimal point if f if and only if:

$$\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq 0,\forall x\in C$$

direction 1:

assume x^* is an optimal point of f , and let's prove:

$$\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq 0,\forall x\in C$$

since f is convex, then from the gradient inequality:

$$f\left(x\right)+\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq f\left(x^*\right)$$

and since x^* is optimal:

$$f\left(x^*\right)\leq f\left(x\right),\forall x\in C$$

hence:

$$f\left(x\right)+\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq f\left(x^*\right)\leq f\left(x\right)$$

$$\boxed{\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq 0,\forall x\in C}$$

direction 2:

assume:

$$\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq 0,\forall x\in C$$

let's prove x^* is an optimal point.

notice that

$$\forall r>0:\nabla f\left(x\right)^{T}\left(x^*-x\right)\leq 0,\forall x\in C\cap B\left(x^*,r\right)$$

$$\forall r>0:f'\left(x,x^*-x\right)\leq 0,\forall x\in C\cap B\left(x^*,r\right)$$

and this is true even when $r \longrightarrow 0$, meaning that all the points around x^*have a decent direction towards x^* no matter how close we get.

It has to mean that every direction from x^* is an accent direction:

$$f'\left(x^*, x - x^*\right) \geq 0, \forall x \in C$$

$$\nabla f\left(x^*\right)^T \left(x - x^*\right) \geq 0, \forall x \in C$$

$$\boxed{\text{hence } x^* \text{ is a stationery point, and since the function is convex, it must be an optimal point.}}$$

b)

Consider the problem:

$$\min\left\{\sum_{i=1}^{n} g\left(x_i\right)\right\}$$
$$s.t : x \in \Delta_n$$

where g is a one dimensional continues function.

i)

Prove that if x^* is an optimal solution of the problem then any permutation of x^* is also an optimal solution

solution:

assume x^* is an optimal solution of the problem, and let y be a permutation of the vector x^*, notice that:

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i^* = 1$$
$$\forall i \in \{1, 2, ..., n\} : y_i \geq 0$$

hence:

$$y \in \Delta_n$$

in addition:

$$\sum_{i=1}^{n} g\left(y_i\right) = \sum_{i=1}^{n} g\left(x_i^*\right)$$

$\boxed{\text{hence } y \text{ is a feasible solution, and it also achieves the same function value as } x^* \text{, hence } y \text{ is also an optimal solution of the problem.}}$

ii)

Show that if g is convex then the vector:

$$y = \left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$$

is an optimal solution of the problem.

solution:

since g is convex, then from Jensen inequality:

$$\sum_{i=1}^{n} \frac{1}{n} g\left(x_i\right) \geq g\left(\sum_{i=1}^{n} \frac{1}{n} x_i\right)$$

hence:

$$\sum_{i=1}^{n} g\left(x_i\right) = n \sum_{i=1}^{n} \frac{1}{n} g\left(x_i\right) \geq ng\left(\sum_{i=1}^{n} \frac{1}{n} x_i\right) = ng\left(\frac{1}{n} \underbrace{\sum_{i=1}^{n} x_i}_{=1}\right) = ng\left(\frac{1}{n}\right)$$

so we have found a lower bound for the objective function, in addition for $x = y$:

$$\sum_{i=1}^{n} g\left(y_i\right) = \sum_{i=1}^{n} g\left(\frac{1}{n}\right) = g\left(\frac{1}{n}\right) \sum_{i=1}^{n} 1 = ng\left(\frac{1}{n}\right)$$

$\boxed{\text{thus the lower bound is attained, hence } y \text{ is an optimal solution of the problem.}}$

Question 2:

Consider the following optimization problem:

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|_2^2 + \|Lx\|_1$$

a)

Show that any vector in $x \in \mathbb{R}^n$ can be written as $x = y + z$ where $y \in Null\,(A) \cap Null\,(L)$ and $z \in Null\,(A) \cap Null\,(L)_{\perp}$

solution:

intuitive, if we pick:

$$y = P_C\,(x)$$

where

$$C = Null\,(A) \cap Null\,(L)$$

and define:

$$z = x - y = x - P_C\,(x)$$

then we expect z to be orthogonal to any vector in C, however we have no theorem that says so, so let's try to prove it.

let's find the projection onto this space, it is equivalent to solving the problem:

$$\begin{aligned} &\min_{y} \|x-y\|_2^2 \\ &s.t: Ay = 0 \\ &\quad\;\; Ly = 0 \end{aligned}$$

this is a convex problem, since the objective function is convex and the constraints are linear. slater condition also holds, for example $y = 0$ is a feasible solution, thus:

$$\{K.K.T\} = \{Optimal\}$$

let's find the K.K.T points, the Lagrangian:

$$L\,(y, \mu_1, \mu_2) = \|x-y\|_2^2 + \mu_1^T\,(Ay) + \mu_2^T Ly$$

the K.K.T conditions:

$$\begin{cases} \frac{\partial L(y,\mu_1,\mu_2)}{y} = -2\left(x-y\right)+A^T\mu_1+L^T\mu_2=0 & (1) \\ Ay=0 & (2) \\ Ly=0 & (3) \end{cases}$$

from (1):

$$-2\left(x-y\right)+A^T\mu_1+L^T\mu_2=0$$

$$-2x+2y+A^T\mu_1+L^T\mu_2=0$$

$$2y=2x-A^T\mu_1-L^T\mu_2$$

$$y=x-\frac{1}{2}\left(A^T\mu_1+L^T\mu_2\right)$$

notice that $C_1=Null\left(A\right)$ is a convex set

take $z_1,z_2\in Null\left(A\right)$ and $\lambda\in[0,1]$ then:

$$A\left(\lambda z_1+\left(1-\lambda\right)z_2\right)=\lambda Az_1+\left(1-\lambda\right)Az_2=0+0=0$$

so:

$$\lambda z_1+\left(1-\lambda\right)z_2\in Null\left(A\right)$$

so $C=Null\left(A\right)\cap Null\left(L\right)$ is a convex set as an intersection between two convex sets.

C is also closed as an intersection between two closed sets.

thus from the first projection theorem we know that $P_C\left(x\right)$ exists and it's unique, hence it doesn't really important to find the Lagrange multipliers, as they can be any two vectors in $\mathbb{R}^n$, the unique solution will still always have the form:

$$y=P_C\left(X\right)=x-\frac{1}{2}\left(A^T\mu_1+L^T\mu_2\right)$$

notice that defining:

$$z=x-y=x-x+\frac{1}{2}\left(A^T\mu_1+L^T\mu_2\right)=\frac{1}{2}\left(A^T\mu_1+L^T\mu_2\right)$$

and if $u\in Null\left(A\right)\cap Null\left(L\right)$ then:

$$z^Tu=\frac{1}{2}\left(A^T\mu_1+L^T\mu_2\right)^Tu=\frac{1}{2}\mu_1^TAu+\frac{1}{2}\mu_2^TLu=0$$

hence indeed:

$$z \in Null\left(A\right) \cap Null\left(L\right)_{\perp}$$

so by defining:

$$y = P_{Null(A) \cap Null(L)}\left(x\right) \in Null\left(A\right) \cap Null\left(L\right)$$

$$z = x - y \in Null\left(A\right) \cap Null\left(L\right)_{\perp}$$

we get:

$$x = y + z$$

b)

Show that this problem is coercive if and only if $Null\left(A\right) \cap Null\left(L\right) = \{0\}$

direction 1:

assume the problem is coercive, let's show that $Null\left(A\right) \cap Null\left(L\right) = \{0\}$.

assume by contradiction that $Null\left(A\right) \cap Null\left(L\right) \neq \{0\}$, then:

$$\exists v \neq 0 \in Null\left(A\right) \cap Null\left(L\right)$$

$$\begin{aligned} f\left(\alpha v\right) &= \left\|A\alpha v - b\right\|_2^2 + \left\|L\alpha v\right\|_1 = \left\|\alpha Av - b\right\|_2^2 + \left\|\alpha Lv\right\|_1 = \\ &= \left\|-b\right\|_2^2 + \left\|0\right\|_1 = \left\|b\right\|_2^2 \xrightarrow{\alpha \to \infty} \left\|b\right\|_2^2 \end{aligned}$$

hence we have found a direction in which the function doesn't go to ∞ as the norm of the vector goes to ∞ which is a contradiction to the coerciveness of the problem.

hence $Null\left(A\right) \cap Null\left(L\right) = \{0\}$.

direction 2:

assume that $Null\left(A\right) \cap Null\left(L\right) = \{0\}$, let's prove the problem is coercive.

if $\forall x : Lx \neq 0 \Rightarrow \left\|Lx\right\|_1 \xrightarrow{\|x\| \longrightarrow \infty} \infty$

if $\forall x : Ax \neq 0 \Rightarrow \left\|Ax - b\right\|_2^2 \xrightarrow{\|x\| \longrightarrow \infty} \infty$

if $\forall x : Ax \neq 0, or, Lx \neq 0 \Rightarrow \max\left\{\left\|Lx\right\|_1, \left\|Ax - b\right\|_2^2\right\} \xrightarrow{\|x\| \longrightarrow \infty} \infty$

$$f\left(x\right)=\left\Vert Ax-b\right\Vert _{2}^{2}+\left\Vert Lx\right\Vert _{1}\geq\max\left\{ \left\Vert Lx\right\Vert _{1},\left\Vert Ax-b\right\Vert _{2}^{2}\right\}$$

and since $Null\left(A\right)\cap Null\left(L\right)=\left\{ 0\right\}$ as we said $\max\left\{ \left\Vert Lx\right\Vert _{1},\left\Vert Ax-b\right\Vert _{2}^{2}\right\} \xrightarrow{\left\Vert x\right\Vert \longrightarrow\infty}\infty$

thus $f\left(x\right)$ is a coercive function.

another way to prove it:

assume by contradiction that $f\left(x\right)$ is not coercive, then there exists some path

$$x\left(t\right)=\left(\begin{array}{c}
f_{1}\left(t\right)\\
f_{2}\left(t\right)\\
.\\
.\\
.\\
f_{n}\left(t\right)
\end{array}\right)$$

in which $\left\Vert x\left(t\right)\right\Vert \xrightarrow{t\longrightarrow\infty}\infty$ but $f\left(x\right)$ doesn't go to infinity, specifically it means that there exist M that is an upper bound on the function on this path.

$$\forall t:\left\Vert f\left(x\left(t\right)\right)\right\Vert \leq M$$

$$\forall t:\left\Vert \left\Vert Ax\left(t\right)-b\right\Vert _{2}^{2}+\left\Vert Lx\left(t\right)\right\Vert _{1}\right\Vert \leq M$$

since $\left\Vert x\left(t\right)\right\Vert \xrightarrow{t\longrightarrow\infty}\infty$ it can only happen if:

$$\exists t_{0}:\forall t>t_{0}:Ax\left(t\right)=0\cap Lx\left(t\right)=0$$

which is a contradiction to the fact that:

$$Null\left(A\right)\cap Null\left(L\right)=\left\{ 0\right\}$$

Question 3:

Let $A \in \mathbb{R}^{nxm}$. Prove that one and only one of the following is true:

(I) $\exists x > 0$ such that $Ax \leq b$ and $Bx < c$

(II) $\exists u, v \geq 0$ such that $b^T u + c^T w \leq 0$,$A^T u + B^T w \geq 0$ and at least one of the following holds:

$$(i)\, w \neq 0$$

$$(ii)\, A^T u + B^T w \neq 0$$

$$(iii)\, b^T u + c^T w < 0$$

direction 1:

Assume that (I) is satisfied, we will prove that (II) is not satisfied.

$$\exists x > 0 : \begin{cases} Ax \leq b \\ Bx < c \end{cases}$$

assume by contradiction that (II) is also satisfied, then:

$$\exists u, w \geq 0 : \begin{cases} b^T u + c^T w \leq 0 \\ A^T u + B^T w \geq 0 \end{cases}$$

and at least one of (i) (ii) and (iii) are satisfied.

in general:

$$A^T u + B^T w \geq 0$$

$$u^T A + w^T B \geq 0$$

$$u^T Ax + w^T Bx \geq 0$$

if $w \neq 0$

$$0 \leq \underbrace{u^T}_{\geq 0} \underbrace{Ax}_{\leq b} + w^T Bx \leq u^T b + \underbrace{w^T}_{>0} \underbrace{Bx}_{<c} \boxed{<} u^T b + w^T c = b^T u + c^T w \leq 0$$

which is a contradiction.

if $A^T u + B^T w \neq 0$ then:

$$u^T Ax + w^T Bx > 0$$

$$0 \boxed{<} \underbrace{u^T}_{\geq 0} \underbrace{Ax}_{\leq b} + w^T Bx \leq u^T b + \underbrace{w^T}_{\geq 0} \underbrace{Bx}_{<c} \leq u^T b + w^T c = b^T u + c^T w \leq 0$$

which is a contradiction.

if $b^T u + c^T w < 0$ then:

$$0 \leq \underbrace{u^T}_{\geq 0} \underbrace{Ax}_{\leq b} + w^T Bx \leq u^T b + \underbrace{w^T}_{\geq 0} \underbrace{Bx}_{<c} \leq u^T b + w^T c = b^T u + c^T w \boxed{<} 0$$

which is a contradiction.

direction 2:

Assume that (I) is not satisfied, we will prove that (II) is satisfied.

We have proved something similar in the homework, Motzkin lemma, and we use the second formulation of Farkas lemma in order to do so. Now we can follow the same proof, using the non homogeneous Farkas lemma

let's remember the non homogeneous Farkas lemma: suppose that there exists $y_0 \geq 0$ such that $A^T y_0 = c$ then exactly one of the following is feasible:

$$(a)\, Ax \leq b, c^T x > d$$

$$(b)\, A^T y = c, b^T y \leq d, y \geq 0$$

which is equivalent to saying the both the following claims are equivalent:

$$(a)\, Ax \leq b \Rightarrow c^T x \leq d$$

$$(b)\, \exists y \in \mathbb{R}^n_+ : A^T y = c, b^T y \leq d$$

(I) is not satisfied, hence the system:

$$\begin{cases} Ax \leq b \\ Bx < c \end{cases}$$

doesn't have a solution, which also means:

$$Ax \leq b \Rightarrow Bx \geq c$$

looking at some row b_j of B as a column vector:

$$Ax \leq b \Rightarrow b_j^T x \geq c_j$$

$$Ax \leq b \Rightarrow -b_j^T x \leq -c_j$$

using the second formulation of the non homogeneous Farkas lemma:

$$\exists y \in \mathbb{R}_+^n : A^T y = -b_j, b^T y \leq -c_j$$

choosing:

$$w = e_j \geq 0, u = y \geq 0$$

then:

$$A^T u + B^T w = A^T y + B^T e_j = -b_j + b_j = 0 \geq 0$$

$$b^T u + c^T w = b^T y + c^T e_j = b^T y + c_j \leq -c_j + c_j = 0 \leq 0$$

and also :

$$w = e_j \neq 0$$

thus (II) is satisfied.

Question 4:

Let $Q \in R^{nxn}$ be a PSD matrix, $b \in Img\left(Q\right)$ and let:

$$f\left(x\right) = \sqrt{x^T Qx - 2b^T x + c}$$

a)

Prove that there exists a matrix $A \in \mathbb{R}^{kxn}$ which is full row rank such that $Q = A^T A$ where k is the rank of Q.

solution:

since Q is a symmetric matrix we can use spectral decomposition:

$$Q = U^T DU$$

since Q is P.S.D, then all it's eigen values are non negative, and we can write:

$$Q = U^T D^{\frac{1}{2}} D^{\frac{1}{2}} U = \left(D^{\frac{1}{2}} U\right)^T \left(D^{\frac{1}{2}} U\right)$$

let's assume, that the eigen values of Q are arranged on D from largest at $D_{1,1}$ to smallest at $D_{n,n}$. if the rank of Q is $k \leq n$, then the matrix $D^{\frac{1}{2}}U$ has $n-k$ rows which are all zeros (the last rows), thus if we just remove those rows, we will get a new matrix A, which is of size kxn and of rank k, hence full row rank, and also:

$$\left(D^{\frac{1}{2}} U\right)^T \left(D^{\frac{1}{2}} U\right) = A^T A = Q$$

thus:

$$\boxed{A = \left(D^{\frac{1}{2}} U\right)_{1:k,:} = \tilde{D}^{\frac{1}{2}} U}$$

b)

Let $z = A\tilde{x}$, for some $\tilde{x} \in \mathbb{R}^n$ where A is the matrix defined in (a). Show that x satisfies $Ax = z$ if and only if there exists a vector $y \in Null\left(Q\right)$ such that $x = \tilde{x} + y$.

direction 1:

assume that $Ax = z$, we will prove that

$$\exists y \in Null\left(Q\right) : x = \tilde{x} + y$$

solution:

$$z = Ax = A\tilde{x}$$

$$Ax = A\tilde{x}$$

$$A\left(x - \tilde{x}\right) = 0$$

$$A^T A\left(x - \tilde{x}\right) = 0$$

$$Q\left(x - \tilde{x}\right) = 0$$

define:

$$y = x - \tilde{x}$$

then:

$$y \in Null\left(Q\right)$$

and:

$$\tilde{x} + y = \tilde{x} + x - \tilde{x} = x$$

direction 2:

assume that

$$\exists y \in Null\left(Q\right) : x = \tilde{x} + y$$

we will prove that $Ax = z$

solution:

$$Qy = 0$$

$$A^T Ay = 0$$

$$y^T A^T Ay = 0$$

$$\left(Ay\right)^T\left(Ay\right) = 0$$

$$\|Ay\|_2^2 = 0$$

$$Ay = 0$$

plugging this:

$$Ax = A(\tilde{x} + y) = A\tilde{x} + Ay = A\tilde{x} = z$$

c)

Show that the domain of f is equivalent to a set of the form:

$$dom f = \left\{x \in \mathbb{R}^n : \|Ax - g(A,b)\|^2 \geq \|g(A,b)\|^2 - c\right\}$$

Where g is a function of A and b, and find an explicit formulation of g.

solution:

the domain of f is the set of all point which satisfies:

$$x^T Q x - 2b^T x + c \geq 0$$

$$x^T A^T A x - 2b^T x + c \geq 0$$

$$\|Ax\|_2^2 - 2b^T x + c \geq 0$$

which is close to what we want but not exactly.

let's try to find $g(A,b)$ from the final form:

$$\|Ax - g(A,b)\|^2 \geq \|g(A,b)\|^2 - c$$

$$x^T A^T A x - 2g(A,b)^T Ax + g(A,b)^T g(A,b) \geq \|g(A,b)\|^2 - c$$

$$\|Ax\|_2^2 - 2g(A,b)^T Ax + \|g(A,b)\|^2 \geq \|g(A,b)\|^2 - c$$

$$\|Ax\|_2^2 - 2g(A,b)^T Ax \geq -c$$

so we want:

$$g(A,b)^T A = b^T$$

$$A^T g(A,b) = b$$

since $b \in Img(Q)$ there:

$$\exists v_b : Qv_b = b$$

define:

$$g(A,b) = Av_b$$

and then:

$$A^T g(A,b) = A^T A v_b = Qv_b = b$$

sanity check:

if we plug $g(A,b) = Av_b$ then:

$$-2g(A,b)^T Ax = -2(Av_b)^T Ax = -2v_b^T A^T Ax = -2b^T x$$

to conclude, by defining:

$$g(A,b) = Av_b$$

where:

$$Qv_b = b$$

we get that the domain of f is of the form:

$$dom f = \left\{x \in \mathbb{R}^n : \|Ax - g(A,b)\|^2 \geq \|g(A,b)\|^2 - c\right\}$$

d)

Find necessary and sufficient conditions on the problem parameters for $f(x)$ to be a convex function.

solution:

Since the $\sqrt{x}$ function is a concave function, the only option for the function to be convex is if it is a norm, meaning:

$$f(x) = \sqrt{x^T Qx - 2b^T x + c}$$

(I'm actually not sure this statement is true but I have no other direction) for this to happen we need that:

$$x^T Qx - 2b^T x + c = \|Bx + d\|_2^2 + e = x^T B^T Bx + 2d^T Bx + d^T d + e$$

so we need:

$$B^T B = Q$$

but we already found such a matrix, $A^T A = Q$, so plug $B = A$.

we also need that:

$$2d^T B = -2b^T$$

$$B^T d = -b$$

plugging $B = A$:

$$A^T d = -b$$

hence we need $b \in Img\left(A^T\right)$ but since $b \in Img\left(Q\right)$ then also $b \in Img\left(A^T\right)$:

$$Qv_b = b$$

$$A^T\left(Av_b\right) = b$$

so we can choose $d = Av_b = g\left(A, b\right)$.

the only thing left to check is:

$$e = c - d^T d \geq 0$$

$$c \geq d^T d = \|d\|_2^2 = \|g\left(A, b\right)\|_2^2$$

thus the function is convex if and only if:

$$c \geq \|g\left(A, b\right)\|_2^2$$

which also implies the the domain won't be empty since:

$$\|g\left(A, b\right)\|^2 - c \leq 0 \leq \|Ax - g\left(A, b\right)\|^2$$

Question 5:

Consider the set $C = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} p_i x_i^2 \leq \alpha\}$ where $p \in \mathbb{R}^n_{++}$

a)

Consider the orthogonal projection problem over the set C:

i)

Formulate the problem. Are KKT conditions necessary for optimality? are they sufficient?

solution:

the problem of finding $P_C(x)$ can be written as:

$$\min_y \|x-y\|_2^2$$
$$s.t: \sum_{i=1}^{n} p_i y_i^2 \leq \alpha$$

The squared norm $f(x) = \|x\|_2^2$ is a convex function, thus the objective function is convex as a affine transformation of the variables of a convex function.

notice that:

$$\sum_{i=1}^{n} p_i y_i^2 = y^T \begin{pmatrix} p_1 & & & & \\ & p_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot & \\ & & & & & p_n \end{pmatrix} y = y^T Q y$$

since $p \in \mathbb{R}^n_{++}$ Q is a strictly diagonal dominant matrix with non negative entries on the diagonal, hence P.D. Therefore $y^T Q y$ is a quadratic function with a P.D matrix, thus convex. Hence the second constraint is a convex set as a level set of a convex function.

Since both the objective function and the constraints are convex, the K.K.T condition are sufficient for optimiality.

if also slater condition holds, then the K.K.T condition are also necessary.

if $\alpha > 0$ then slater condition holds, for example for $y = 0$

$$\sum_{i=1}^{n} p_i y_i^2 = 0 < \alpha$$

if $\alpha \leq 0$ we cannot satisfy slater condition (actually if $\alpha < 0$ the problem is not feasible and if $\alpha = 0$ only $y = 0$ is feasible) .

to sum:
for every α the K.K.T condition are sufficient.
for $\alpha > 0$ they are also necessary.

ii)

Use the K.K.T conditions in order to devise a simple algorithm that finds an optimal solution of the problem

solution:

The lagrangian:

$$L(y, \lambda) = \|x - y\|_2^2 + \lambda \left(y^T Q y - \alpha \right)$$

let's write the K.K.T conditions.

$$\begin{cases} \frac{\partial L(y,\lambda)}{\partial y} = -2(x - y) + 2\lambda Q y = 0 & (1) \\ \lambda \left(y^T Q y - \alpha \right) = 0 & (2) \\ y^T Q y \leq \alpha & (3) \\ \lambda \geq 0 & (4) \end{cases}$$

if $\lambda = 0$, then from (1):

$$-2(x - y) = 0$$

$$y = x$$

but we also need to satisfy (3) :

$$y^T Q y = x^T Q x \leq \alpha$$

which is not always true, this is the case where x is already in C.

if $\lambda > 0$ then from (1):

$$-2(x-y) + 2\lambda Q y = 0$$

$$-2x + 2y + 2\lambda Q y = 0$$

$$2y + 2\lambda Q y = 2x$$

$$(I + \lambda Q) y = x$$

since Q is P.D then also $I + \lambda Q$ is P.D and invertible.

$$y = (I + \lambda Q)^{-1} x$$

from (2):

$$y^T Q y - \alpha = 0$$

$$\left((I + \lambda Q)^{-1} x\right)^T Q \left((I + \lambda Q)^{-1} x\right) - \alpha = 0$$

$$x^T (I + \lambda Q)^{-1} Q (I + \lambda Q)^{-1} x - \alpha = 0$$

$$\sum_{i=1}^{n} \left(\frac{p_i}{(1 + \lambda p_i)^2} x_i^2\right) - \alpha = 0$$

notice that when $\lambda \longrightarrow 0$ then:

$$\sum_{i=1}^{n} \left(\frac{p_i}{(1 + \lambda p_i)^2} x_i^2\right) \longrightarrow \sum_{i=1}^{n} \left(p_i x_i^2\right) > \alpha$$

because if $\sum_{i=1}^{n} \left(p_i x_i^2\right) \leq \alpha$, we are in the case when no projection is needed, hence when $\lambda \longrightarrow 0$:

$$\phi(\lambda) = \sum_{i=1}^{n} \left(\frac{p_i}{(1 + \lambda p_i)^2} x_i^2\right) - \alpha > 0$$

also when $\lambda \longrightarrow \infty$:

$$\phi(\lambda) = \sum_{i=1}^{n} \left(\frac{p_i}{(1 + \lambda p_i)^2} x_i^2\right) - \alpha = -\alpha < 0$$

(we assume we are in the feasible non trivial case where $\alpha > 0$)

thus $\phi(\lambda)$ has a root in $(0, \infty)$, also it's a decreasing function in λ thus we can find it's root using the bisection algorithm.

to conclude, the algorithm will be:

if $\alpha < 0$: return there is no feasible solution.

if $\alpha = 0$ return $y = 0$

if $\alpha > 0$:

if $(\sum_{i=1}^{n} p_i x_i^2 \leq \alpha)$ return:

$$y = x$$

else return:

$$y = (I + \lambda Q)^{-1} x$$

where λ is the only root of the function:

$$\phi(\lambda) = \sum_{i=1}^{n} \left(\frac{p_i}{(1 + \lambda p_i)^2} x_i^2 \right) - \alpha$$

in $(0, \infty)$

b)

Let y be a given vector. Consider the problem (P) where we aim to find the furthest point from y which is contained in C.

i)

Formulate problem (P) as an optimization problem. is it convex?

solution:

we can write P as:

$$\max_{x} \|y - x\|_2^2$$
$$s.t: \sum_{i=1}^{n} p_i x_i^2 \leq \alpha$$

or:

$$\min_{x} - \|y - x\|_2^2$$
$$s.t: \sum_{i=1}^{n} p_i x_i^2 \leq \alpha$$

from the same reasons of section a, the objective function and constraint are a convex, however, now we are maximizing a convex function over a convex set, which is identical to minimizing a concave function over a convex set. hence this is not a convex optimization problem.

ii)

Compute the dual of problem (P) as a single variable problem. Is it convex?

solution:

we will compute the dual of:

$$\min_{x} -\|y-x\|_2^2$$
$$s.t: \sum_{i=1}^{n} p_i x_i^2 \leq \alpha$$

The Lagrangian:

$$L(x,\lambda) = -\|y-x\|_2^2 + \lambda\left(x^T Q x - \alpha\right)$$

$$\begin{aligned}
q(\lambda) = \min_x L(x,\lambda) &= \min_x -\|y-x\|_2^2 + \lambda\left(x^T Q x - \alpha\right) = \\
&= \min_x -x^T x + 2y^T x - y^T y + \lambda x^T Q x - \alpha\lambda = \\
&= \min_x x^T(\lambda Q - I)x + 2y^T x - y^T y - \alpha\lambda = \\
&= \min_x \sum_{i=1}^{n}\left((\lambda p_i - 1)x_i^2 + 2y_i x_i\right) - y^T y - \alpha\lambda = \\
&= \sum_{i=1}^{n}\min_{x_i}\left((\lambda p_i - 1)x_i^2 + 2y_i x_i\right) - y^T y - \alpha\lambda
\end{aligned}$$

we need to find a minimum to the 1D function:

$$\min_{x_i}\left((\lambda p_i - 1)x_i^2 + 2y_i x_i\right)$$

if $\lambda p_i - 1 > 0$ this is a simple parabola with a minimum, the minimum is attained at:

$$x_i^* = -\frac{2y_i}{2(\lambda p_i - 1)} = \frac{y_i}{1-\lambda p_i}$$

and the minimal value is:

$$\begin{aligned} A &= (\lambda p_i - 1)\left(\frac{y_i}{1-\lambda p_i}\right)^2 + 2y_i\left(\frac{y_i}{1-\lambda p_i}\right) = \\ &= -(1-\lambda p_i)\cdot\frac{y_i^2}{(1-\lambda p_i)^2} + \frac{2y_i^2}{1-\lambda p_i} = \\ &= -\frac{y_i^2}{(1-\lambda p_i)} + \frac{2y_i^2}{1-\lambda p_i} = \frac{y_i^2}{1-\lambda p_i} \end{aligned}$$

if $\lambda p_i - 1 = 0$, then this function becomes a linear function and only has a minimum if $y_i = 0$

if $\lambda p_i - 1 < 0$ this is a parabola that doesn't have a minimum value hence:

$$\min_{x_i}\left((\lambda p_i - 1)\, x_i^2 + 2y_i x_i\right) = \begin{cases} \frac{y_i^2}{1-\lambda p_i} & \lambda p_i - 1 > 0 \\ 0 & \lambda p_i - 1 = 0 \cap y_i = 0 \\ -\infty & \lambda p_i - 1 < 0 \cup (\lambda p_i - 1 = 0 \cap y_i \neq 0) \end{cases}$$

notice that if $\lambda p_i - 1 \geq 0$:

$$\lambda p_i \geq 1$$

$$\lambda \geq \frac{1}{p_i} > 0$$

$$\lambda \geq \max_i\left\{\frac{1}{p_i}\right\}$$

$$\lambda \geq \frac{1}{\min_i p_i}$$

hence the dual problem is:

$$\begin{aligned} &\max_{\lambda} \sum_{i:\lambda p_i - 1 > 0} \frac{y_i^2}{1-\lambda p_i} - \|y\|_2^2 - \alpha\lambda \\ &s.t : \lambda \geq \frac{1}{\min_i p_i} \end{aligned}$$

with the small addition that λ can't be equal to $\frac{1}{\min_i p_i}$ if $y_{\arg\min_i p_i} = 0$.

The dual problem is always a convex problem, in the sense that it maximizes a concave function over a convex set.

iii)

Show that the optimal primal solution is attained and strong duality holds.

solution:

if we can find a pair x, λ which are both feasible and such that:

$$x \in \arg\min_{x} L(x, \lambda)$$

$$\lambda \left(\sum_{i=1}^{n} p_i x_i^2 - \alpha \right) = 0$$

that this pair are optimal solution to the dual and primal problems respectively, and strong duality is satisfied.

since any feasible λ satisfies

$$\lambda \geq \frac{1}{\min_i p_i} > 0$$

we have to look for some x that satisfies

$$\sum_{i=1}^{n} p_i x_i^2 = \alpha$$

let's show that we can find some $\lambda > \frac{1}{\min_i p_i}$,which is a feasible solution to the dual problem, and also satisfies:

$$\sum_{i=1}^{n} p_i x_i^2 = \alpha$$

since $\lambda > \frac{1}{\min_i p_i}$ if we define x in the following way:

$$x_i = \frac{y_i}{1 - \lambda p_i}$$

then:

$$x \in \arg\min_{x} L(x, \lambda)$$

let's find this λ:

$$\sum_{i=1}^{n} p_i x_i^2 = \sum_{i=1}^{n} p_i \left(\frac{y_i}{1 - \lambda p_i} \right) = \sum_{i=1}^{n} p_i \frac{y_i^2}{(1 - \lambda p_i)^2} = \alpha$$

define:

$$\phi(\lambda) = \sum_{i=1}^{n} p_i \frac{y_i^2}{(1 - \lambda p_i)^2} - \alpha$$

when $\lambda \longrightarrow \frac{1}{\min_i p_i}$ then $\phi(\lambda) \longrightarrow \infty$

when $\lambda \longrightarrow \infty$ then $\phi(\lambda) \longrightarrow -\alpha < 0$

hence this system has a solution λ^* in $\left(\frac{1}{\min_i p_i}, \infty \right)$.

hence the pair λ^* and x^* defined as:

$$x_i^* = \frac{y_i}{1 - \lambda^* p_i}$$

satisfies:

$$x^* \in \arg\min_x L(x, \lambda)$$

$$\lambda^* \left(\sum_{i=1}^{n} p_i x_i^{*2} - \alpha \right) = 0$$

thus they are optimal solutions for the dual and primal problem, and strong duality holds.

c)

Consider the problem of finding the minimum radius ball which contains set C

i)

Formulate the problem as an optimization problem on the center of the ball y and the radius r.

solution:

We can write the problem as follows:

$$\begin{aligned} &\min_{y,r} r \\ &s.t: \|x - y\|_2^2 \le r^2, \forall x \in \{C\} \end{aligned}$$

ii)

Use the dual formulation in (b) to reformulate the problem as:

$$\begin{aligned} &\min_{y,\rho,\lambda} \rho \\ &s.t: \sum_{i:p_i\lambda-1>0} y_i^2 \left(\frac{p_i\lambda}{p_i\lambda - 1} \right) + \alpha\lambda \le \rho \\ &\quad \lambda \ge \frac{1}{p^*} \end{aligned}$$

where $p^* = \min_i p_i$

solution:

we need to formulate our problem a little bit differently.

remember that in section b we solved the problem $\max_{x\in C}\|x-y\|_2^2$.

let's define for each y the function $P_{\bar{C}}(y)=\arg\max_{x\in C}\|x-y\|_2^2$.

we know the radios we are looking for will satisfy:

$$\|x-y\|_2^2\leq\rho,\forall x\in C$$

if and only if the center will be close enough to the furthest point from it in C meaning (y is the center):

$$\|P_{\bar{C}}(y)-y\|_2^2\leq\rho$$

hence we can formulate a new problem:

$$\begin{aligned}&\min_{y,r}\rho\\&s.t:\|P_{\bar{C}}(y)-y\|_2^2\leq\rho\end{aligned}$$

since we have shown that strong duality holds we can find an explicit solution for $\|P_{\bar{C}}(y)-y\|_2^2$.

we know that the the the pair $P_{\bar{C}}(y)$ and λ that solve the primal and dual problem have to satisfy:

$$\begin{aligned}\|P_{\bar{C}}(y)-y\|_2^2&=\max_{x\in C}\|x-y\|_2^2=-\min_{x\in C}-\|x-y\|_2^2=-q(\lambda)=\\&=-\sum_{i:\lambda p_i-1>0}\frac{y_i^2}{1-\lambda p_i}+\|y\|_2^2+\alpha\lambda=\\&=\sum_{i:\lambda p_i-1>0}\left(\frac{y_i^2}{(\lambda p_i-1)}+y_i^2\right)+\alpha\lambda=\\&=\sum_{i:\lambda p_i-1>0}\left(\frac{y_i^2+y_i^2\lambda p_i-y_i^2}{\lambda p_i-1}\right)+\alpha\lambda=\\&=\sum_{i:\lambda p_i-1>0}\left(\frac{y_i^2\lambda p_i}{\lambda p_i-1}\right)+\alpha\lambda\end{aligned}$$

where we also need to remember that λ has to be in the appropriate domain, hence the new problem can be formulated as:

$$\begin{aligned}
&\min_{y,\rho,\lambda} \rho \\
&s.t: \sum_{i:\lambda p_i - 1>0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1}\right) + \alpha\lambda \leq \rho \\
&\lambda \geq \frac{1}{\min_i p_i}
\end{aligned}$$

iii)

Show that (P') is jointly convex in y, ρ, λ

solution:

the objective function is linear hence convex.

$$\sum_{i:\lambda p_i - 1>0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1}\right) = y^T \begin{pmatrix} \frac{\lambda p_1}{\lambda p_1 - 1} & & & & & \\ & \frac{\lambda p_2}{\lambda p_2 - 1} & & & & \\ & & . & & & \\ & & & . & & \\ & & & & . & \\ & & & & & \frac{\lambda p_n}{\lambda p_n - 1} \end{pmatrix} y = y^T R y$$

since $\lambda \geq \frac{1}{\min_i p_i}$ all the entries in the diagonal of R are non negative, hence R is a P.D matrix, thus $y^T R y$ is convex function as a quadratic function with a P.D matrix(and under linear transformation of the variables y, ρ, λ)

$\alpha\lambda - \rho$ is a convex function as a linear function of the variables.

hence the constraint is a convex set as a level set of a convex function.

The second constraint is linear hence convex.

the objective function is convex and the constraint defines a convex set as an intersection of convex sets, thus this is a convex optimization problem in λ, ρ, y

iv)

Solve problem (P')

solution:

I'm actually not sure how to solve it. we can try using the K.K.T conditions but it will be difficult to solve.

The Lagrangian:

$$L\left(\rho, y, \lambda, \eta_1, \eta_2\right) = \rho + \eta_1 \left(\sum_{i:\lambda p_i - 1 > 0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1} \right) + \alpha\lambda - \rho \right) + \eta_2 \left(\frac{1}{\min_i p_i} - \lambda \right)$$

the K.K.T conditions:

$$\begin{cases} \begin{cases} \frac{\partial L(\rho,y,\lambda,\eta_1,\eta_2)}{\partial \rho} = 1 - \eta_1 = 0 & (1) \\ \frac{\partial L(\rho,y,\lambda,\eta_1,\eta_2)}{\partial \lambda} = \eta_1 \sum_{i:\lambda p_i - 1 > 0} \frac{y_i^2 p_i (\lambda p_i - 1) - y_i^2 \lambda p_i^2}{(\lambda p_i - 1)^2} + \eta_1 \alpha - \eta_2 = 0 & (2) \\ \frac{\partial L(\rho,y,\lambda,\eta_1,\eta_2)}{\partial y_i} = \eta_1 \forall i : \lambda p_i - 1 > 0 : \eta_1 \frac{2 y_i \lambda p_i}{\lambda p_i - 1} = 0 & (3) \end{cases} \\ \eta_1 \left(\sum_{i:\lambda p_i - 1 > 0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1} \right) + \alpha\lambda - \rho \right) = 0 & (4) \\ \eta_2 \left(\lambda - \frac{1}{\min_i p_i} \right) = 0 & (5) \\ \sum_{i:\lambda p_i - 1 > 0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1} \right) + \alpha\lambda \le \rho & (6) \\ \lambda \ge \frac{1}{\min_i p_i} & (7) \\ \eta_1, \eta_2 \ge 0 & (8) \end{cases}$$

from (1):

$$\eta_1 = 1 \ge 0$$

from (3):

$$\eta_1 \frac{2 y_i \lambda p_i}{\lambda p_i - 1} = 0$$

$$\forall i : \lambda p_i - 1 > 0 : y_i = 0$$

from (2):

$$\begin{aligned} &\eta_1 \sum_{i:\lambda p_i - 1 > 0} \frac{y_i^2 p_i \left(\lambda p_i - 1\right) - y_i^2 \lambda p_i^2}{\left(\lambda p_i - 1\right)^2} + \eta_1 \alpha - \eta_2 = \\ &= \eta_1 \sum_{i:\lambda p_i - 1 > 0} \frac{y_i^2 \lambda p_i^2 - y_i^2 p_i - y_i^2 \lambda p_i^2}{\left(\lambda p_i - 1\right)^2} + \eta_1 \alpha - \eta_2 = \\ &= \eta_1 \sum_{i:\lambda p_i - 1 > 0} \frac{-y_i^2 p_i}{\left(\lambda p_i - 1\right)^2} + \alpha + \eta_2 = \alpha - \eta_2 = 0 \end{aligned}$$

$$\eta_2 = \alpha \geq 0$$

from (4):

$$\sum_{i:\lambda p_i - 1 > 0} \left(\frac{y_i^2 \lambda p_i}{\lambda p_i - 1} \right) + \alpha\lambda - \rho = 0$$

$$\alpha\lambda - \rho = 0$$

$$\alpha\lambda = \rho$$

$$\lambda = \frac{\rho}{\alpha}$$

from (5):

$$\lambda - \frac{1}{\min_i p_i} = 0$$

$$\lambda = \frac{1}{\min_i p_i}$$

$$\rho = \frac{\alpha}{\min_i p_i}$$

since the problem is convex with convex constraints, K.K.T condition are sufficient for optimiality, hence the solution is:

$$\rho = \frac{\alpha}{\min_i p_i}$$

$$y = 0$$