SSY281 Model Predictive Control

Assignment 3 - Optimization basics and QP problems

Fengxiang Xue(LoginID:xuefe)

February 21, 2024

Question 1 Constrained optimization

(a) A strictly convex function is ,for any two distinct points x_1 and x_2 in its domain and for any θ between 0 and 1 has the following inequality:

$$f(\theta x_1 + (1-\theta)x_2) < \theta f(x_1) + (1-\theta)f(x_2)$$

this means that the function value at any point on the straight line segment joining x_1 and x_2 is strictly less than the weighted average of the function values at x_1 and x_2

(b) A convex set in a vector space is a set that, for every pair of points within it, contains the whole line segment that connects them. This means if you take any two points in a convex set and draw a line between them, every point on that line will also be within the set.

A set S $\subseteq$R^n is convex if for every x_1 and $x_2$$\in S, and for every$
θ between 0 and 1 the following holds:

$$\theta x_1 + (1-\theta)x_2 \in S$$

(c) An optimization problem is considered a convex optimization problem if it satisfies the following conditions: 1. the objective function f is convex 2. the intersection of convex sets is a convex set; g_i(x) is convex 3. The equality function h is affine:

$$h_j(x) = A_j x - b_j$$

where A_j is matrix and b_j is a vector.
When these conditions are met, the feasible region defined by the constraints g(x) $\leqslant$ 0 and h(x) =0 is a convex set, and the objective f(x) is a convex over this set.

Question 2 Convexity

(a) S_1 is convex. To prove S_1

$$S_1 = \{x \in R^n | \alpha \leq \alpha^T x \leq \beta\}$$

is convex, we need to show that for any two points x, y $\in$S$_1$ and any λ that 0$\leqslant\lambda \leqslant 1, the point z =$$\lambda$x +(1-$\lambda$)y also lies in S_1:

$$z = \lambda x + (1-\lambda)y \in S, 0 \leq \lambda \leq 1$$

let x,y be any 2 points in S_1,by definition of S_1 we have

$$\alpha \leqslant \alpha^T x \leqslant \beta, \alpha \leqslant \alpha^T y \leqslant \beta$$

consider a point z that is a convex :

$$z = \lambda x + (1-\lambda)y, 0 \leqslant \lambda \leqslant 1$$

$$\alpha^T z = \alpha^T(\lambda x + (1-\lambda)y) = \lambda\alpha^T x + (1-\lambda)\alpha^T y;$$

$$\alpha \leq \lambda\alpha^T x \leq \beta; \alpha \leq \lambda\alpha^T y \leq \beta;$$

and

$$0 \leq \lambda \leq 1; 0 \leq 1-\lambda \leq 1$$

so the combination of above also lies in 0 to 1, which is:

$$\alpha \leq \lambda\alpha^T x + (1-\lambda)\alpha^T y \leq \beta$$

This shows that z satisfies the condition to be in S_1, and thus S_1 is convex.

(b) S_2 is convex.

$$S_2 = \{x | \parallel x - y \parallel \leq f(y), \forall y \in S\}$$

where

$$S \subseteq R^n, f(y) \geq 0, \forall y,$$

and $\parallel . \parallel$is an arbitrary norm.

we take x_1 and x_2 in S_2 and 0$\leq\lambda \leq 1$:

$$\parallel \lambda x_1 + (1-\lambda)x_2 - y \parallel \leq f(y)$$

for all y $\in S$.

1. by definition of S_2,

$$\parallel x_1 - y \parallel \leq f(y) and \parallel x_2 - y \parallel \leq f(y)$$

for all y $\in S$,

2. consider z for:

$$z = \lambda x_1 + (1-\lambda)x_2$$

then

$$\| z - y \| = \| \lambda x_1 + (1-\lambda)x_2 - y \|$$

3.Using triangle inequality and the distribution property of scalar multiplication over addition:

$$\| z - y \| = \| \lambda x_1 + (1-\lambda)x_2 - \lambda y - (1-\lambda)y \|$$

$= \| \lambda(x_1 - y) + (1-\lambda)(x_2 - y) \|$
$\leq \lambda \underbrace{(\| x_1 - y \|)}_{\leq f(y)} + (1-\lambda) \underbrace{(\| x_2 - y \|)}_{\leq f(y)}$

4. so $\| z - y \| \leq f(y), \forall y \in S, and z \in$S$_2$. This shows S_2 is convex because the convex combination of any two points in S_2 is also contained in S_2 .

(c) S_3 is not convex.

$$S_3 = \{(x,y) | y \leq 2^x, \forall (x,y) \in R^2\}$$

consider two arbitrary points(x_1,y_1) and (x_2,y_2) must also satisfy y$\leq 2^x$, which means:

$$y_1 \leq 2^{x_1}; y_2 \leq 2^{x_2}$$

for S_3 to be convex, every point (x,y) on the segment between (x_1,y_1) and (x_2,y_2) must also satisfy y $\leq 2^x$. A point (x,y) on the line segment can be written as:

$$(x,y) = \lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) = \underbrace{\lambda x_1 + (1-\lambda)x_2}_{x} + \underbrace{\lambda y_1 + (1-\lambda)y_2}_{y}$$

for some $\lambda \in [0,1]$.

according to definition of S_3,

$$y \leq 2^x \Rightarrow \underbrace{\lambda y_1 + (1-\lambda)y_2}_{y} \leq 2^{\overbrace{\lambda x_1 + (1-\lambda)x_2}^{x}} \leq \lambda 2^{x_1} + (1-\lambda)2^{x_2};$$

which is not generally true, because of non-linearity of the exponential function, therefore S_3 is not convex.

To give a concrete counterexample:
consider two points (0,1) , (1,2) , and their middle point(1/2, 3/2); (0,1) and (1,2) both lie in S_3 because of $1 \leq 2^0$ and $2 \leq 2^1$, however the midpoint(1/2, 3/2) doesn't satisfy the S_3 inequality:

$$\frac{3}{2} \nleq 2^{\frac{1}{2}}$$

Question 3 Norm problems as linear programs

(a) The equation(3) is given by: $min_x \parallel Ax - b \parallel_\infty$;

This means we want to find the x that minimize $\parallel Ax - b \parallel_\infty$;

the equation(5) seeks:

$$min_{x,\epsilon}\epsilon, s.t. -\epsilon \leq (Ax-b)_i \leq \epsilon, \forall i$$

for problem 5,we want to find the maximum x that the ith ϵ can satisfy the inequality equation,

$$\parallel (Ax-b)_i \parallel \leq \epsilon, \forall i \in 1, ...n$$

These 2 equations are both aiming to find the smallest x value that satisfies the constraints for all elements of Ax-b.

(b) equation(2) is :

$$min_z C^T z, s.t.Fz \leq g,$$

Assuming

$$z^T = [x^T \epsilon]$$

if x is of size n, then

$$c^T = [0, 0, ..., 0, 1]$$

because we are minimizing ϵ, the cost vector c will have zeros for all elements corresponding to x and 1 for the element corresponding to ϵ. this is we are not directly minimizing any component of x but rather ϵ.

to find F and g that conclude

$$-\epsilon \leq (Ax-b)_i \leq \epsilon$$

which means

$$(Ax)_i - \epsilon \leq b_i$$
$$-(Ax)_i - \epsilon \leq -b_i$$

we can write $Fx \leq g$ as:

$$\underbrace{\begin{bmatrix} a_1^T & -1 \\ a_2^T & -1 \\ ... & ... \\ a_n^T & -1 \\ -a_1^T & -1 \\ -a_2^T & -1 \\ ... & ... \\ -a_n^T & -1 \end{bmatrix}}_{F} * \begin{bmatrix} x_i \\ \epsilon \end{bmatrix} \leq \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_n \\ -b_1 \\ -b_2 \\ ... \\ -b_n \end{bmatrix}}_{g}$$

$$g = \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_n \\ -b_1 \\ -b_2 \\ ... \\ -b_n \end{bmatrix}$$

or can be rewritten as:

$$\underbrace{\begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} x \\ \epsilon \end{bmatrix}}_{z} \leq \underbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}_{g}$$

(c) optimal epsilon= 0.4583;
optimal x=

$$\begin{bmatrix} -2.0674 \\ -1.1067 \end{bmatrix}$$

(d) we have linear programs:

$$min_z c^T z$$

$$s.t. Fz \leq g$$

For each constraint $Fz - g \leq 0$, we introduce a dual variable,$y \geq 0$.
The dual problem thus becomes:

$$maximize : -g^T y;$$

$$s.t. F^T y + c = 0; y \geq 0$$

Lagrangian

$$L(z, y) = c^T z + y^T (Fz - g)$$

(e) To solve the y in dual problem, we use linprog() function, with the definition of 'f' = g; Aeq = F transpose; beq = -c;
we get the optimal Lagrange multipliers

$$y = \begin{bmatrix} 0 \\ 0 \\ 0.4095 \\ 0.4284 \\ 0 \\ 0.1621 \\ 0 \\ 0 \end{bmatrix}$$

(f) stationarity condition for the KKT is given by the Lagrangian gradient to z

$$\nabla_z L(z, y) = c + F^T y = 0;$$

we can find the optimal z by linprog() function with previous Lagrangian multiplier 'y', inequality parameter F and g , and we know c,

$$z = \begin{bmatrix} x \\ \epsilon \end{bmatrix} = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix}$$

which is the same with previous primal solution.

Question 4 Quadratic programming

(a) With quadprog() function we can solve the x1,x2,u0,u1 as:

$$\begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 0.49996 \\ 1.9 \\ -0.5 \end{bmatrix}$$

objective function minimization f(x,u) = 10.36

(b) Yes,KKT hold at the solution in (a), active lower bound constraints for x_1, active upper bound constraints for x_2.

To check if the KKT conditions hold, we are looking for:

1.Primal feasibility :x_1, x_2, u_0, u_1 within these bounds, the primal feasibility condition is satisfied.

2.Dual Feasibility: This condition requires that all Lagrange multipliers for inequality constraints be non-negative.

lambda.lower= $\begin{bmatrix} 9.2000 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

lambda.upperr= 1.0e-03 * $\begin{bmatrix} 0 \\ 0.1471 \\ 0 \\ 0 \end{bmatrix}$

lambda.eqlin= $\begin{bmatrix} 3.8000 \\ -1.0001 \end{bmatrix}$

(lambda.ineqlin = 0×1 empty double column vector:

since we don't have Ax+b$\leq$ 0 such constraint here.)

3.Stationarity:Since quadprog has returned an optimal solution, we can assume it has found a point where the gradient of the Lagrangian is zero , satisfying the stationarity condition.

4.Complementary Slackness:This condition states that the product of each Lagrange multiplier and the corresponding constraint should be zero. This condition also was satisfied, if we try to use lambda.lower and lambda.upper and lambda.eqlin to multiply with x, we found they are all equal to 0.

(c) 1.Removing the lower bound on x_1: since now x_1 is already on its lower bound,removing this constraint could allow x_1 to decrease further, expanding x_1's feasible area.

2.Removing the upper bound on x_1: removing the upper bound may not affect the optimal solution since x_1 is not approaching to its upper bound.