AI2101/EE5606 Convex Optimization

Lecture 9

Scribe(s): AI20BTECH11006, Chirag Mehta

April 14, 2022

Quadratic Problems

A convex quadratic program is just a least squares problem with linear constraints. Consider the following formulation

$$\min_{\underline{x}} \ \underline{x}^T Q \underline{x} + \underline{c}^T \underline{x} \tag{1}$$

$$\text{s.t: } C\underline{x} \leq \underline{d} \tag{2}$$

Recall: Every positive definite matrix is a gram matrix

We can write

$$\exists R \ : \ Q = R^T R \tag{3}$$

We can rewrite our optimization problem as

$$\min_{\underline{x}} \underline{x}^T R^T R \underline{x} + \underline{c}^T \underline{x} \tag{4}$$

$$\text{s.t: } C\underline{x} \leq \underline{d} \tag{5}$$

Further, we can write

$$\exists \underline{b} \ : \ \underline{c} = 2R^T \underline{b} \tag{6}$$

[1] Now, we can rewrite the optimization problem as a least square problems using (4) and (6)

[1] since R^T is a full rank matrix, a unique solution must exist for the above system of equations.

$$\min_{\underline{x}} \ \|R\underline{x} - \underline{b}\|_2^2 - \underline{b}^T \underline{b} \tag{7}$$

$$\text{s.t: } C\underline{x} \leq \underline{d} \tag{8}$$

$$\underline{c} = 2R^T \underline{b}$$

Problem: Given two polyhedra, determine the distance between them.

$$P_1 : \ A\underline{x} \leq \underline{b} \tag{9}$$

$$P_2 : \ C\underline{x} \leq \underline{d} \tag{10}$$

The optimization problem can be formulated as follows

$$\min_{\underline{x_1}, \underline{x_2}} \ \|\underline{x_1} - \underline{x_2}\|_2 \tag{11}$$

$$\text{s.t: } \underline{x_1} \in P_1 \tag{12}$$

$$\underline{x_2} \in P_2$$

The above is a quadratic problem with linear constraints, next we will look at a Quadratically Constrained Quadratic Programs (QCQP)

$$\min_{\underline{x}} \underline{x}^T Q \underline{x} + \underline{b}^T \underline{x} \tag{13}$$

$$\text{s.t: } \underline{x}^T Q_i \underline{x} + \underline{b_i}^T \underline{x} \leq c_i \tag{14}$$

$$C\underline{x} \leq \underline{d}$$

Example: Portfolio optimization

$$\min_{\underline{x}} \ \underline{x}^T \Sigma \underline{x} - \lambda \underline{\mu}^T \underline{x} \tag{15}$$

$$\text{s.t: } \underline{x}^T \underline{1} = 1 \tag{16}$$

$$\underline{x} \geq 0$$

where Σ is the covariance matrix that accounts for the risk. [2] The above problem is a quadratic program since the objective is quadratic, while the constraints are linear.

[2] if λ is small then we are taking low risk, λ controls the risk against expected returns.

Example: Linear Discriminator Lets say we have a dataset of two classes

$$C1: \ \{x_1, x2, \ldots, x_m\}$$

$$C2: \ \{y_1, y2, \ldots, y_m\}$$

Check whether the given datapoints are linearly separable.

Initial Thought: Check if the convex hulls of the given datapoints intersect, if yes, then the classes are not linearly separable.[3]

[3] Note that this problem is a feasibility problem.

Another method: Consider the following two inequalities for respective classes

$$C1: \ \underline{a}^T \underline{x_i} - b \geq 1 \tag{17}$$

$$C2: \ \underline{a}^T \underline{y_i} - b \leq -1$$

We can frame our feasibility problem as

$$\min_{\underline{a}, b} \ 0 \tag{18}$$

$$\text{s.t: } \underline{a}^T \underline{x_i} - b \geq 1 \ , \ i = 1, 2, \ldots, m \tag{19}$$

$$\underline{a}^T \underline{y_i} - b \leq -1 \ , \ i = 1, 2, \ldots, m$$

[4] Both the listed methods are equivalent.

[4] The reason why we choose 1 instead of 0 in the inequality is because of inequalities are treated the same way as equalities is cvxpy solver, the problem will then always be feasible by choose a as zero vector, and b to be 0

Further, we can extend the problem from a feasibility problem to hard margin support vector machine.

$$\min_{\underline{a}, b} \ \|\underline{a}\|_2^2 \tag{20}$$

$$\text{s.t: } \underline{a}^T \underline{x_i} - b \geq 1 \ , \ i = 1, 2, \ldots, m \tag{21}$$

$$\underline{a}^T \underline{y_i} - b \leq -1 \ , \ i = 1, 2, \ldots, m$$

This is a quadratic program (QP), we can however, reformulate this problem as follows

$$\min_{\underline{a},b,t} \ t \tag{22}$$

$$\text{s.t: } \underline{a}^T \underline{x_i} - b \geq t \ , \ i = 1,2,\ldots,m \tag{23}$$

$$\underline{a}^T \underline{y_i} - b \leq -t \ , \ i = 1,2,\ldots,m$$

$$\|\underline{a}\|_2 \leq 1$$

The above formulation in Quadratically Constrained Quadratic Program (QCQP). [5]

[5] Since the margin in our latter problem is $\frac{2t}{\|a\|_2}$, but we are maximizing t, then we would get $\|a\|_2 = 1$

Assertion: The above two problems are equivalent. Solving one would also solve the other, we can arrive at solution for latter from the prior formulation

Lets say after solving the prior, we get $\underline{a}'$, b' as the optimal solution, we can divide both $\underline{a}'$, b' by $\|\underline{a}'\|_2$, to get the solution to our latter formulation.

Second Order Cone Prgrams

A second order cone can be represented as follows

$$\{(\underline{x}, t) : \ \|\underline{x}\|_2 \leq t\} \tag{24}$$

We can take another example

$$\{(x, y) : \ xy \geq 1\} \tag{25}$$

The above can be formulated as a cone in the following manner

$$\{(x,y) : \left\| \begin{pmatrix} x - y \\ 2 \end{pmatrix} \right\|_2 \geq x + y\} \tag{26}$$

A general form of SOCP is given below

$$\min_{\underline{x}} \ \underline{e}^T \underline{x} \tag{27}$$

$$\text{s.t: } \|A_i \underline{x} + \underline{b_i}\|_2 \leq \underline{c_i}^T \underline{x} + d_i \ , \ i = 1,2,\ldots,m \tag{28}$$

$$F\underline{x} = \underline{g}$$

Consider the following problem

$$\max_{\underline{x}} \ \left(\frac{1}{\underline{a_1}^T \underline{x} + b_1} + \frac{1}{\underline{a_2}^T \underline{x} + b_2} + \cdots + \frac{1}{\underline{a_m}^T \underline{x} + b_m} \right)^{-1} \tag{29}$$

We can reformulate this problem as a SOCP in the following manner

$$\min_{\underline{t}} \ t_1 + t_2 + \cdots + t_m \tag{30}$$

$$\text{s.t: } t_i \left(\underline{a_i}^T \underline{x} + b_i \right) \geq 1 \ , \ i = 1,2,\ldots,m \tag{31}$$

$$t_i \geq 0 \ , \ i = 1,2,\ldots,m$$

[6] The above problem is now a SOCP.

Now, lets look at yet another problem which can be formulated as a SOCP

$$\max_{\underline{y}} \quad (y_1, y_2)^{0.5} \tag{32}$$
$$\text{s.t: } \underline{y} = A\underline{x} + \underline{b} \tag{33}$$
$$\underline{y} \geq 0$$

We can use the hypograh trick to convert this to a SOCP

$$\max_{\underline{y},t} \quad t \tag{34}$$
$$\text{s.t: } \underline{y} = A\underline{x} + \underline{b} \tag{35}$$
$$\underline{y} \geq 0$$
$$\left\| \begin{pmatrix} y_1 - y_2 \\ 2t \end{pmatrix} \right\|_2 \leq y_1 + y_2$$
$$t \geq 0$$

We can generalise the above problem to a more broader problem, consider the following problem with $\underline{y}$ being a 4 dimensional vector this time

$$\max_{y_1, y_2} \quad (y_1 y_2 y_3 y_4)^{0.25} \tag{36}$$
$$\text{s.t: } \underline{y} = A\underline{x} + \underline{b} \tag{37}$$
$$\underline{y} \geq 0$$

Now, take $t_1^2 \leq y_1 y_2$, $t_2^2 \leq y_3 y_4$, $t^2 \leq t_1 t_2$. We can formulate the problem as

$$\max_{\underline{y}, t_1, t_2, t} \quad t \tag{38}$$
$$\text{s.t: } \underline{y} = A\underline{x} + \underline{b} \tag{39}$$
$$y_1 y_2 \geq t_1^2$$
$$y_3 y_4 \geq t_2^2$$
$$t_1 t_2 \geq t^2$$
$$t \geq 0, \ \underline{y} \geq 0$$
$$t_1 \geq 0, \ t_2 \geq 0$$

Further, using (26), we can represent the constraints as conic constraints. The above problem can be generalised further to $\underline{y}$ being a 2^n dimensional vector.

[6] We want $\underline{a_i}^T \underline{x} + b_i$ to be positive, so we impose non-negative condition on t to implicitly use this constraint

Robust Linear Programming

Consider the following problem

$$\min_{\underline{x}} \ \underline{c}^T \underline{x} \tag{40}$$

$$\text{s.t: } \underline{a_i}^T \underline{x} \leq b_i, \ i = 1, 2, \ldots, m \tag{41}$$

the catch is that there is uncertainity in $\underline{a_i}$

We can take the following case

$$\underline{a_i} \sim \mathcal{N}(\underline{\mu_i}, \Sigma_i) \tag{42}$$

We want $\underline{a_i}^T \underline{x} \leq b_i$ to hold with a high probability, say

$$\mathrm{P}(\underline{a_i}^T \underline{x} \leq b_i) \geq \eta \tag{43}$$

[7] Now call $\underline{a_i}^T \underline{x}$ as u_i, we know that

[7] where η is some number close to 1 like 0.99.

$$\mathrm{E}[u_i] = \underline{\mu_i}^T \underline{x} \tag{44}$$

$$var(u_i) = \sigma^2 = \underline{x}^T \Sigma_i \underline{x} \tag{45}$$

Now, we can rewrite our problem as

$$\mathrm{P}\left(\frac{u_i - \mathrm{E}[u_i]}{\sigma} \leq \frac{b_i - \mathrm{E}[u_i]}{\sigma}\right) \tag{46}$$

[8] Using Z score, we get

[8] We used this transformation to write the equation in terms of Z score

$$\frac{b_i - \mathrm{E}[u_i]}{\sigma} \geq Q^{-1}(\eta) \tag{47}$$

[9] using (45)

[9] Here Q^{-1} is a function that maps probabilities to Z scores

$$b_i - \mathrm{E}[u_i] \geq Q^{-1}(\eta) \left(\underline{x}^T \Sigma_i \underline{x}\right)^{0.5} \tag{48}$$

$$b_i - \mathrm{E}[u_i] \geq Q^{-1}(\eta) \|\Sigma \underline{x}\|_2^{0.5} \tag{49}$$

the above problem is now a SOCP.