1(a)

09 May 2019 09:44

This is done in the lecture. That proof uses a 1-d characterization. An alternate easier way to prove is to show that the quadratic function $f(x) = 0.5x'Px+q'x+r$ is convex and then C will be convex simply because C is the level set of f at 0 and by theorem : "All (non-empty) level sets of convex function are convex". f is convex can be easily showed using Hessian (which is P) is psd (P is given psd). Theorem 21.0.5.

If you did not solve this part it shows you did not even revise lectures!

1(b)

09 May 2019 09:54

Counter example: Consider P is negative identity matrix, $q=0,r=0$. Then $C=\{x | -\|x\|^2 \le0\}$= all vectors, which is obviously convex. Therefore converse is NOT true.

1(c)

09 May 2019 09:56

It is enough if you roughly draw an ellipse.

If you drew anything other than ellipse, perhaps you must repeat class X or IX, where conic sections are taught!

2(a)

09 May 2019 09:59

This was derived in lecture and shown that the subdifferential set is the q-norm unit ball, where $1/p+1/q=1$.

2(b)

09 May 2019 10:00

This is a nice quadratic, (hence) differentiable function. We computed gradient of this in the lecture (and it is a very common example otherwise too): x-a. Hence gradient at 0 is -a. Thus sub-differential set is \{-a\}

If you did not do this in exam, then you must repeat class XI etc, where (partial) derivative of quadratic is taught.

2(c)

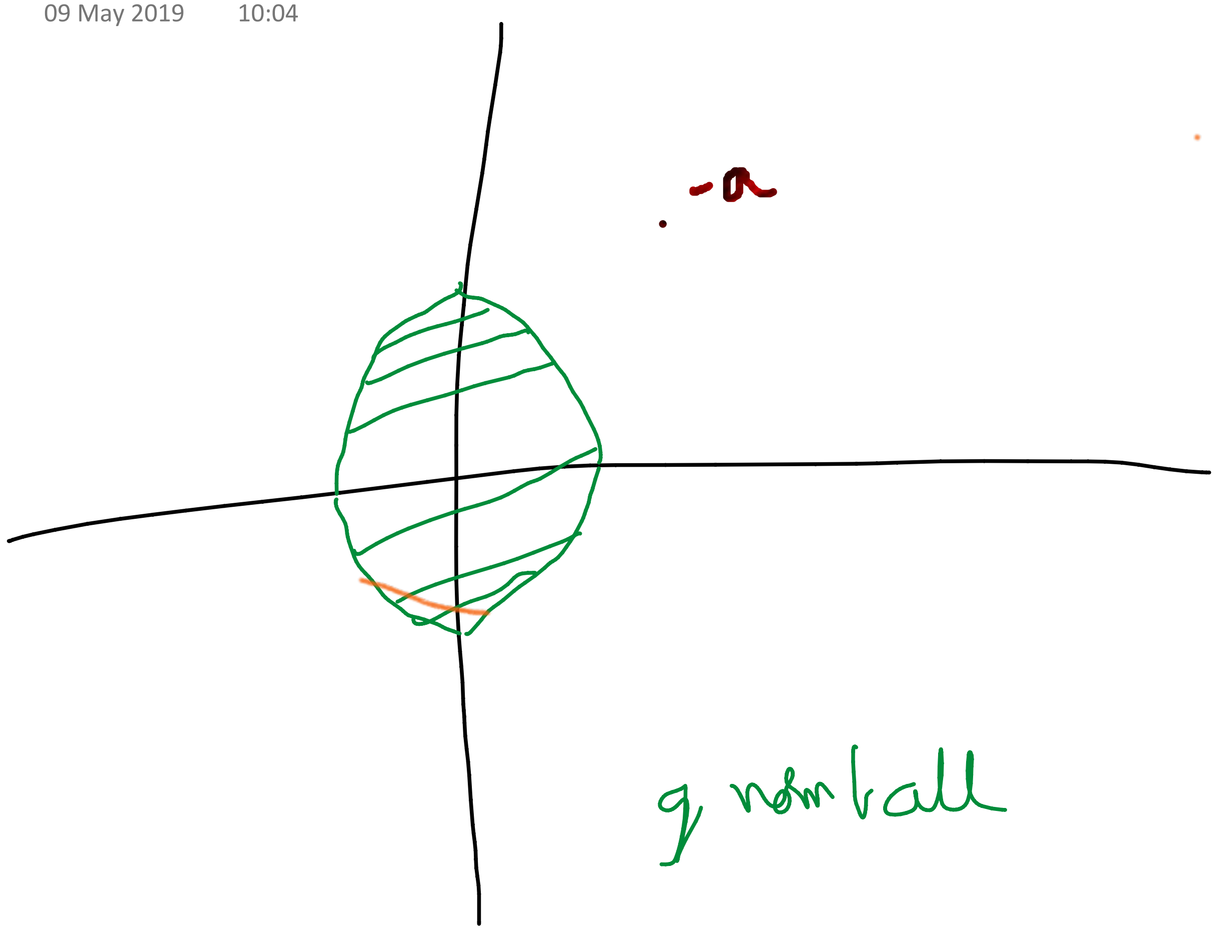
09 May 2019 10:03

Using theorem 21.0.1, the sub-differential set is convex hull of the subdifferential sets of the functions that are active. If $a \ne 0$, then only h is active at zero, hence the set is $\{0\}$. Else ($a=0$), then both are active but subdiff of h is contained in that of g, hence this set is q-norm unit ball.

If you did not do this in exam, it shows you do not remember key results from lectures/notes.

2(d)

09 May 2019 10:04



MPs

03 May 2019 15:33

Let 'x' represent desired location to be found

Let $u_1, \dots, u_m$ denote the coordinates of the settlements.

CP1

$$\min_{x \in \mathbb{R}^n} \max_{i \in \{1, \dots, m\}} \|x - u_i\|$$

$n=2$

CP2

$$\min_{x \in \mathbb{R}^n, t \in \mathbb{R}} t$$

s.t.

$$\|x - u_i\| \le t \quad \forall i = 1 \text{ to } m$$

$$\|x - u_i\| \le t \iff t^2 - \|x - u_i\|^2 \ge 0, t \ge 0$$

$\Updownarrow$ Claim I

$$\begin{bmatrix} t I_n & x - u_i \\ (x - u_i)^T & t \end{bmatrix} \succeq 0$$

Proof of claim I: If $t = 0$,

LHS $\iff x = u_i$

RHS $\iff x = u_i$

Let $t \neq 0$, then LHS $\iff t > 0$, $t^2 - \|x - u_i\|^2 \ge 0$

RHS $\iff t > 0$, $t - \frac{\|x - u_i\|^2}{t} \ge 0$

by Schur complement lemma.

SDP

03 May 2019 16:12

CP3

$$\min_{z \in \mathbb{R}^{n+1}} \; c^T z$$

s.t. $$B - \sum_{i=1}^{n+1} A_i z_i \succeq 0$$

I_n is identity of size n.

$\mathbb{1}_i$ is vector of zeros and 1 at i^{th} position.

O_n is square matrix of size n all zeros

$$c = \begin{bmatrix} 0_{n\times 1} \\ 1 \end{bmatrix} \qquad B = \begin{bmatrix} \begin{bmatrix} O_n & -\mu_1 \\ -\mu_1^T & 0 \end{bmatrix} & & O \\ & \ddots & \\ O & & \begin{bmatrix} O_n & -\mu_m \\ -\mu_m^T & 0 \end{bmatrix} \end{bmatrix}$$

$$A_i = \begin{bmatrix} \begin{bmatrix} O_n & \mathbb{1}_i \\ \mathbb{1}_i^T & 0 \end{bmatrix} & & O \\ & \ddots & \\ O & & \begin{bmatrix} O_n & \mathbb{1}_i \\ \mathbb{1}_i^T & 0 \end{bmatrix} \end{bmatrix}$$

$i = 1$ to n

$$A_{n+1} = \begin{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & 1 \end{bmatrix} & & O \\ & \ddots & \\ O & & \begin{bmatrix} I_n & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix}$$

CP4

03 May 2019 16:18

$$\min_{x \in \mathbb{R}^n} \max_{i \in \{1,\dots,m\}} \|x - \mu_i\|^2$$

$$= \min_{x \in \mathbb{R}^n,\, t \in \mathbb{R}} t$$

$$\underline{\underline{\text{s.t.}}} \quad \|x - \mu_i\|^2 - t \leq 0$$

$\Big\}$ $\underline{\underline{CP4}}$

CP5

03 May 2019 16:20

$$\mathcal{L}(x,t,\lambda) \equiv t + \sum_{i=1}^{m} \lambda_i \left(\|x-\mu_i\|^2 - t \right)$$

$$\underline{\mathcal{L}}(\lambda) \equiv \min_{x\in\mathbb{R}^n,\, t\in\mathbb{R}} \left(1-\sum_{i=1}^{m}\lambda_i\right) t + \sum_{i=1}^{m} \lambda_i \|x-\mu_i\|^2$$

$$= \begin{cases} -\infty & \sum_{i=1}^{n}\lambda_i \neq 1 \\ \min_{x\in\mathbb{R}^n} \sum_{i=1}^{n}\lambda_i\|x-\mu_i\|^2 & \sum_{i=1}^{n}\lambda_i = 1 \end{cases}$$

$$x^* \xrightarrow{\text{solves}} \iff 2\sum_{i=1}^{n}\lambda_i(x^*-\mu_i) = 0 \qquad (\text{gradient} = 0)$$

$$\iff x^* = \sum_{i=1}^{n}\lambda_i\mu_i$$

$$\therefore \min_{x\in\mathbb{R}^n} \sum_{i=1}^{n}\lambda_i\|x-\mu_i\|^2 = \sum_{i=1}^{n}\lambda_i \Big\| \sum_{j=1}^{n}\lambda_j\mu_j - \mu_i \Big\|^2 \qquad A = [\mu_1 \cdots \mu_m]$$

$$= \sum_{i=1}^{n}\lambda_i \|A\lambda - \mu_i\|^2$$

$$= \sum_{i=1}^{n}\lambda_i\, \lambda^T A^T A\lambda - 2\underbrace{\sum_{i=1}^{n}\lambda_i\mu_i^T}_{=\lambda^T A^T} A\lambda + \sum_{i=1}^{n}\lambda_i\|\mu_i\|^2$$

$$= \sum_{i=1}^{n}\lambda_i\, \lambda^T A^T A\lambda + \sum_{i=1}^{n}\lambda_i\|\mu_i\|^2$$

$$= \left(\sum_{i=1}^{n}\lambda_i - 2\right)\lambda^T A^T A\lambda + \sum_{i=1}^{n}\lambda_i\|\mu_i\|^2$$

$$= -\lambda^T A^T A\lambda + \ldots$$

linearly concave quadratic!

CP5

$$\max_{\lambda \in \mathbb{R}^m} \; -\lambda^T A^T A \lambda + \sum_{i=1}^{n} \lambda_i \|u_i\|^2$$

s.t. $\lambda \geq 0, \; \sum_{i=1}^{m} \lambda_i = 1$

$\hookrightarrow$ clearly ...

$A^T A \lambda =$

$\sum_i u_i u_i^T \lambda =$

$\|u_1\|^2$

$\vdots$

$\|u_n\|^2$

SDP dual.

03 May 2019 16:36

$$\min_{x\in V} \langle c, x\rangle \quad \text{s.t.} \quad b - \ell(x) \in K \qquad = \qquad \max_{y\in W} -\langle b, y\rangle \quad \underline{\underline{\text{s.t.}}} \quad y \in K^*, \ \ell^T(y) + c = 0.$$

For SDP, V is $\mathbb{R}^n$, W is S_n. K is psd cone $= K^*$.

$$\max_{Y\in S_n} -\operatorname{tr}(BY) \quad \underline{\underline{\text{s.t.}}} \quad Y \succeq 0, \ c_i + \langle A_i, Y\rangle_F = 0 \quad \forall i$$

$$\ell(x) = \sum_{i=1}^{n} x_i A_i$$

$$\langle \ell(x), Y\rangle_F = \left\langle \sum_{i=1}^{n} x_i A_i, Y \right\rangle_F$$

$$= \sum_{i=1}^{n} x_i \langle A_i, Y\rangle_F$$

$$\equiv \ell^T(Y)\, x \qquad \ell^T(Y) \equiv \begin{bmatrix} \langle A_1, Y\rangle_F \\ \vdots \\ \langle A_n, Y\rangle_F \end{bmatrix}$$

CP6

03 May 2019 16:52

Since all B, A_i are block diagonal for CP3, we can safely assume Y is also block diagonal.

Y is symmetric iff all Y_i are symmetric. $Y \succeq 0 \iff Y_i \succeq 0 \ \forall i$

$$B_i = \begin{bmatrix} 0_n & -\mu_i \\ -\mu_i^T & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} Y_1 & & 0 \\ & \ddots & \\ 0 & & Y_m \end{bmatrix}$$

$$\left\langle \begin{bmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{bmatrix}, \begin{bmatrix} Y_1 & & 0 \\ & \ddots & \\ 0 & & Y_m \end{bmatrix} \right\rangle = \sum_{i=1}^m \operatorname{tr}(B_i Y_i) = \sum_{i=1}^m -\operatorname{tr}(\mu_i \gamma_i^T) - \gamma_i^T \mu_i = -2\sum_{i=1}^m \gamma_i^T \mu_i$$

Let $Y_i = \begin{bmatrix} Z_i & \gamma_i \\ \gamma_i^T & \alpha_i \end{bmatrix}$ $\quad$ Z_i are symmetric.

$$\langle A_i, Y \rangle = \sum_{j=1}^m \operatorname{tr}\left(\begin{bmatrix} 0_n & 1_i \\ 1_i^T & 0 \end{bmatrix}, \begin{bmatrix} Z_j & \gamma_j \\ \gamma_j^T & \alpha_j \end{bmatrix} \right) = \sum_{j=1}^m 2\operatorname{tr}(\gamma_j^T 1_i) = 2\sum_{j=1}^m \gamma_{ji}$$

$$\langle A_{n+1}, Y \rangle = \sum_{j=1}^m \operatorname{tr}(Y_j) = \sum_{j=1}^m \operatorname{tr}(Z_j) + \alpha_j$$

$$\max_{\substack{Z_1 \ldots Z_m \in S_n,\ \gamma_1 \ldots \gamma_m \in \mathbb{R}^n, \\ \alpha_1 \ldots \alpha_m \in \mathbb{R}}} -2\sum_{i=1}^m \gamma_i^T \mu_i$$

s.t. $\begin{bmatrix} Z_i & \gamma_i \\ \gamma_i^T & \alpha_i \end{bmatrix} \succeq 0$, $\quad \sum_{j=1}^m \gamma_j = 0$, $\quad \sum_{j=1}^m \operatorname{tr}(Z_j) + \alpha_j = 1$

e.g. as ... is continuous.

s.t. $\begin{bmatrix} Z_i & \gamma_i \\ \gamma_i^T & \alpha_i \end{bmatrix} \succeq 0, \quad \sum_{j=1} \ldots$

$\lambda_i = |\gamma_i|$

w.l.o.g. as objective is continuous

Let's try to eliminate Z_i, α_i.

If $Z_i \succ 0$, then $\begin{bmatrix} Z_i & \gamma_i \\ \gamma_i & \alpha_i \end{bmatrix} \succeq 0 \iff \alpha_i - \gamma_i^T Z_i^{-1} \gamma_i \geq 0$ (∵ Schur complement lemma)

$\iff \gamma_i^T Z_i^{-1} \gamma_i \leq \alpha_i$

Now, $\sum_{i=1}^m \alpha_i = 1 - \sum_{i=1}^m \mathrm{tr}(Z_i)$ and $\alpha_i \geq \gamma_i^T Z_i^{-1} \gamma_i \; \forall i$. So by eliminating α_i's we get:

$$\sum_{i=1}^m \gamma_i^T Z_i^{-1} \gamma_i \leq 1 - \sum_{i=1}^m \mathrm{tr}(Z_i)$$

∴ we are OK as long as $\exists Z_i \succ 0 \ni$

$$\iff \sum_{i=1}^m \min_{Z_i \succ 0} \left[\mathrm{tr}(Z_i) + \gamma_i^T Z_i^{-1} \gamma_i \right] \leq 1$$

$$\min_{Z \succ 0} \mathrm{tr}(Z) + \gamma^T Z^{-1} \gamma = \min_{L \in \mathcal{O}} \min_{\lambda_1, \ldots, \lambda_n > 0} \underbrace{\sum_{i=1}^n \lambda_i + \gamma^T L \Lambda^{-1} L^T \gamma} = \sum_{i=1}^n \min_{\lambda_i > 0} \lambda_i + \frac{\tilde{\gamma}_i^2}{\lambda_i}$$

Let $L^T \gamma = \tilde{\gamma}$

$$= 2 \sum_{i=1}^n |\tilde{\gamma}_i| = 2 \|L^T \gamma\|_1$$

orthogonal matrices

$Z = L \Lambda L^T$

$$= 2 \min_{L \in \mathcal{O}} \|L^T \gamma\|_1 = 2 \|\gamma\|_2$$

(∵ $\|\gamma\|_1 \geq \|\gamma\|_2$ & $\|L\gamma\|_2 = \|\gamma\|_2$)

∴ ... $\max \; -2 \sum_{i}^{m} \gamma_i^T \mu_i$

$$\therefore \quad CP6 = \max_{y_1 \ldots y_m \in \mathbb{R}^n} -2 \sum_{i=1}^{m} y_i^{\top} \mu_i$$

$$\underline{\underline{s.t.}} \quad 2 \sum_{i=1}^{m} \|y_i\|_2 \leq 1 \;, \; \sum_{i=1}^{m} y_i = 0$$

Infact CP2 is an instance of conic-quadratic program, which is a special case of SDP, which is in turn a special case of conic programs. Infact if one writes the conic dual of CP2 (viewing as conic-quadratic program; see (QD) on page50 of Nemirovski book), then one would get exactly same as CP6, however with a lot of ease!

KKT

03 May 2019 18:01

$$\lambda_i^* \left(\|x^* - \mu_i\|^2 - t \right) = 0$$

$$x^* = \sum_i \lambda_i^* \mu_i$$

Conic optimality.

03 May 2019 13:04

$$\langle c, x^* \rangle = \langle -\ell^T(y^*), x^* \rangle$$

$$= -\langle y^*, \ell(x^*) \rangle$$

$$= \langle y^*, b - \ell(x^*) \rangle - \langle y^*, b \rangle$$

$$\therefore \langle y^*, b - \ell(x^*) \rangle = 0 \text{ is the nec. cond.}$$

Note that in the special case of linear programs (K is first orthant), then this simplifies to the well known complementary slackness conditions. So people usually refer to the above as complementary slackness condition for conic programs.