098311 Optimization 1 Spring 2018
HW 9

Chen Tessler 305052680
Orr Krupnik 302629027

June 14, 2018

Solution 1. We can present the problem in an alternative form, which corresponds to the alternative formulation of Farkas' Lemma: Let $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m\times n}$. Then the following two claims are equivalent:

(A) The implication $Ax \le b \Rightarrow c^T x \le d$ holds true.

(B) There exists $y \in \mathbb{R}_+^m$ such that $A^T y = c$ and $b^T y \le d$.

Recall the original (alternative formulation of) Farkas' Lemma:
Let $\tilde{c} \in \mathbb{R}^{n+1}$ and $\tilde{A} \in \mathbb{R}^{m\times(n+1)}$. Then the following two claims are equivalent:

($\tilde{A}$) The implication $\tilde{A}x \le 0 \Rightarrow \tilde{c}^T x \le 0$ holds true.

($\tilde{B}$) There exists $y \in \mathbb{R}_+^m$ such that $\tilde{A}^T y = \tilde{c}$.

and note that now $x \in \mathbb{R}^{n+1}$.

Now we prove our claim.
$(A) \Rightarrow (B)$: Assume some x obeys $Ax \le b$, and that the implication in (A) holds true. Define:

$$\tilde{A} = (A, -b)$$
$$\tilde{c} = (c^T, -d)^T$$

and note that $\tilde{A}, \tilde{c}$ together with $\tilde{x} = (x^T, 1)^T$ implicate exactly that the implication in $(\tilde{A})$ holds true. This means, using Farkas' Lemma, that $(\tilde{B})$ there exists $y \in R_+^m$ such that $\tilde{A}^T y = \tilde{c}$. This, in turn, implies (B) since we have that there exists some $y \in \mathbb{R}_+^n$ such that $A^T y = c$:

$$(\tilde{B})\;\; \tilde{A}^T y = \tilde{c} \iff (A, -b)^T y = (c^T, -d) \iff b^T y = d \Rightarrow b^T y \le d \;\; (B)$$

$(B) \Rightarrow (A)$: Assume (B) holds, i.e. for some $y \in \mathbb{R}_+^n$ for which $A^T y = c$ (which we know exists), we also have $b^T y \le d$. Then, if there exists $x \in \mathbb{R}^m$ such that $Ax \le b$, then:

$$d \le b^T y \le (Ax)^T y = x^T A^T y = x^T c = c^T x$$

Solution 2. $(I) \Rightarrow \neg(II)$: we multiply both sides of the equality in (II) by d^T:

$$0 = d^T A^T u + d^T B^T y = \underbrace{(Ad)^T}_{<0} \underbrace{u}_{\geq 0, \neq 0} + \underbrace{(Bd)^T}_{\leq 0} \underbrace{y}_{\geq 0} < 0$$

Every element of $(Ad)^T$ is negative, and every element of $(Bd)^T$ is non positive. Therefore, since u is non-negative and not the zero vector, the multiplication $(Ad)^T u$ must be strictly negative. In the same way, y is non-negative and therefore $(Bd)^T y$ is non-positive. Together, this contradicts (II) and therefore $(I) \Rightarrow \neg(II)$.

$\neg(II) \Rightarrow (I)$: If (II) doesn't have a solution, it does not have a solution for **any** y, specifically for $y = 0$. Therefore, using Gordan's Alternative Theorem, since we cannot have $A^T u = 0$ for $u \geq 0, u \neq 0$, we have that the system $Ad < 0$ has a solution. Now, we show $Bd \leq 0$. Assume $Bd > 0$, then we can select some $y \neq 0$ and $u \geq 0, u \neq 0$ such that the following is feasible:

$$\begin{aligned} & d^T A^T u + d^T B^T y = 0 \\ \iff & -\underbrace{(Ad)^T u}_{<0} = \underbrace{(Bd)^T y}_{>0} \\ \iff & \sum_{i=1}^{n} (-Ad)_i u_i = \sum_{i=1}^{n} (Bd)_i y_i \end{aligned}$$

specifically, i and j such that $(-Ad)_i > 0$ and $(Bd)_i > 0$: $u = -e_i \cdot -\frac{1}{(Ad)_1}$ and $y = e_j \cdot \frac{1}{(Bd)_1}$ which contradicts the assumption that (II) does not have a solution. Hence we conclude that $\neg(II) \Rightarrow (I)$.

Solution 3.

1. We can write the problem as:

$$\begin{aligned} \min\, & x^T A x \\ \text{s.t. } & a_1^T x \leq b_1 \\ & a_2^T x \leq b_2 \end{aligned}$$

where $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $a_1 = (-1, -1, -3)^T$, $a_2 = (1, 0, 0)T$, $b_1 = -4$ and $b_2 = 2$.
Since the objective function is quadratic with a PD matrix A, it is convex. The constraints are linear. Therefore, this is a convex problem with linear constraints, and we can write the KKT conditions:

$$0 = \nabla f(x^*) + \sum_{i=1}^{2} \lambda_i a_i = \begin{pmatrix} 2x_1^* \\ 4x_2^* \\ 2x_3^* \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ -1 \\ -3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2x_1^* - \lambda_1 + \lambda_2 \\ 4x_2^* - \lambda_1 \\ 2x_3^* - 3\lambda_1 \end{pmatrix}$$

$$0 = \lambda_1 (a_1^T x^* - b_1) = \lambda_1 (-x_1^* - x_2^* - 3x_3^* + 4)$$

$$0 = \lambda_2 (a_2^T x^* - b_2) = \lambda_2 (x_1^* - 2)$$

2. The objective function is a quadratic function with a PD coefficient matrix. Therefore, it is coercive (as we've seen previously in the course) and as such, receives a minimum over the closed constraint set. Since the objective function is also strictly convex, the optimum is unique.
   In a convex problem with linear constraints, KKT conditions are sufficient and necessary. Therefore, the optimal solution must satisfy the KKT conditions.

3. We shall now solve the KKT system:

$$\begin{cases} 2x_1 - \lambda_1 + \lambda_2 = 0 \\ 4x_2 - \lambda_1 = 0 \\ 2x_3 - 3\lambda_1 = 0 \\ \lambda_1(-x_1 - x_2 - 3x_3 + 4) = 0 \\ \lambda_2(x_1 - 2) = 0 \end{cases} \Rightarrow \begin{cases} x_2 = \frac{\lambda_1}{4} \\ x_3 = \frac{3\lambda_1}{2} \\ \lambda_1(-x_1 - \frac{\lambda_1}{4} - 3\frac{3\lambda_1}{2} + 4) = 0 \\ \lambda_2(x_1 - 2) = 0 \end{cases} \Rightarrow$$

$$\begin{cases} \lambda_2 \neq 0, x_1 = 2 : -2\lambda_1 - \frac{\lambda_1^2}{4} - \frac{9\lambda_1^2}{2} + 4\lambda_1 = 0 \Rightarrow \lambda_1 \left(2 - \frac{19}{4}\lambda_1\right) = 0 \\ \Rightarrow \lambda_1 = 0 (\Rightarrow \lambda_2 = 0 \Rightarrow \text{contradiction}) \text{ or } \lambda_1 = \frac{8}{19} \\ \Rightarrow 4 - \frac{8}{19} + \lambda_2 = 0 \Rightarrow \lambda_2 = -\frac{68}{19} < 0 \text{ contradiction} \\ \text{ or} \\ \lambda_2 = 0, x_1 = 0 : \lambda_1 = \lambda_2 = 0 \Rightarrow x_2 = 0, x_3 = 0 \text{ not a feasible solution} \\ \text{ or} \\ \lambda_2 = 0, x_1 \neq 2 : x_1 = \frac{\lambda_1}{2} \Rightarrow -\frac{\lambda_1^2}{2} - \frac{\lambda_1^2}{4} - \frac{9\lambda_1^2}{2} + 4\lambda_1 = 0 \Rightarrow \lambda_1 \left(4 - \frac{21}{4}\lambda_1\right) = 0 \\ \Rightarrow \lambda_1 = 0 (\text{not feasible as seen above}), \text{ or } \lambda_1 = \frac{16}{21} \end{cases}$$

$$\Rightarrow x_1^* = \frac{\lambda_1}{2} = \frac{8}{21}$$

$$x_2^* = \frac{\lambda_1}{4} = \frac{4}{21}$$

$$x_3^* = \frac{3\lambda_1}{2} = \frac{8}{7}$$

Check feasibility where $\lambda_i > 0$:

$$x_1^* = \frac{8}{21} \leq 2$$

Calculating the optimal solution, we get $f(x^*) = 1.524$

Solution 4. 1. The constraint defines a compact set and since the objective is continuous there exists an optimal solution to the problem (Weirstrass).

2.

$$\begin{cases} 2x_1 + 4\lambda x_1^3 = 0 \\ -2x_2 + 4\lambda x_2^3 = 0 \\ -2x_3 + 4\lambda x_3^3 = 0 \\ \lambda(x_1^4 + x_2^4 + x_3^4 - 1) = 0 \end{cases}$$

$$\begin{cases} 2x_1(1+2\lambda x_1^2)=0 \\ 2x_2(-1+2\lambda x_2^2)=0 \\ 2x_3(-1+2\lambda x_3^2)=0 \\ \lambda(x_1^4+x_2^4+x_3^4-1)=0 \end{cases}$$

Note that due to $2x_1(1+2\lambda x_1^2)=0$ and $\lambda \geq 0$ then $x_1 = 0$.
If $\lambda = 0$: $x_1 = x_2 = x_3 = 0 \Rightarrow f(x) = 0$
otherwise ($\lambda > 0$):
If $x_2 = 0$, we have that $x_3 = \pm 1$ ($\lambda = 1$) and when $x_3 = 0$ then $x_2 = \pm 1$ ($\lambda = 1$).
Otherwise when $x_2 \neq 0, x_3 \neq 0 \Rightarrow x_2 = \pm x_3$:

$$2x_2^4 = 1 \Rightarrow x_2 = \pm\sqrt[4]{\frac{1}{2}}$$

Hence the following points are the KKT points:

$$\begin{aligned} &(0, \sqrt[4]{\tfrac{1}{2}}, \sqrt[4]{\tfrac{1}{2}}) \\ &(0, -\sqrt[4]{\tfrac{1}{2}}, \sqrt[4]{\tfrac{1}{2}}) \\ &(0, \sqrt[4]{\tfrac{1}{2}}, -\sqrt[4]{\tfrac{1}{2}}) \\ &(0, -\sqrt[4]{\tfrac{1}{2}}, -\sqrt[4]{\tfrac{1}{2}}) \\ &(0, 1, 0) \\ &(0, -1, 0) \\ &(0, 0, 1) \\ &(0, 0, -1) \end{aligned}$$

3. Since there is only one constraint, its gradient at any point is a linearly independent set, therefore KKT conditions are necessary. In this case at least one of the points above is an optimal solution to the problem (since KKT conditions are neccessary for optimality).
   Due to symmetry this can be reduced to the following three points:

$$\begin{aligned} &(0, \pm\sqrt[4]{\tfrac{1}{2}}, \pm\sqrt[4]{\tfrac{1}{2}}) \Rightarrow f(x) = -\sqrt{2} \\ &(0, \pm 1, 0) \equiv (0, 0, \pm 1) \Rightarrow f(x) = -1 \end{aligned}$$

   thus we conclude that $x^* = (0, \pm\sqrt[4]{\frac{1}{2}}, \pm\sqrt[4]{\frac{1}{2}})$ and $f(x^*) = -\sqrt{2}$.

Solution 5.

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} ||Ax - b||_2^2 \\ &\text{s.t. } e^T x = \alpha \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $e = (1, 1, ..., 1)^T \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ is a parameter.

1. The objective function $||Ax-b||_2^2$ is continuous, coercive and convex (as seen in previous HW & in class), additionally, the constraint defines a closed set and hence the objective attains a minimum over the set.

   As the objective $f(x)$ and constraint $h(x)$ are continuously differentiable functions over $\mathbb{R}^n$ and there is a single constraint (hence it defines a linearly independent set), KKT conditions are necessary for optimality.
   KKT conditions are:

   $$\nabla_x \left(f(x) + \mu h(x)\right) = \nabla_x \left((Ax-b)^T(Ax-b) + \mu(e^T x - \alpha)\right) = 2A^T Ax - 2A^T b + \mu e = 0 \tag{1}$$

   $$e^T x - \alpha = 0$$

   $\Leftarrow$ Direction 1: Assume x, y are optimal solutions to $P(\alpha)$ such that $x \neq y$. Then:

   $$e^T x = e^T y = \alpha \iff e^T(x-y) = 0 \iff (x-y) \in Null(e^T)$$

   Additionally, plugging into the first KKT condition:

   $$\begin{gathered} 2A^T Ax - 2A^T b + \mu e = 0 = 2A^T Ay - 2A^T b + \mu e \\ 2A^T Ax = 2A^T Ay \\ 2A^T A(x-y) = 0 \iff (x-y) \in Null(A) \end{gathered}$$

   Now, since $x-y \in Null(A)$ and $x-y \in Null(e^T)$ then $x-y \in Null(A) \cap Null(e^T)$ and since $x \neq y$, $x-y \neq 0$ and $Null(A) \cap Null(e^T) \neq \{0\}$. Therefore, $Null(A) \cap Null(e^T) = \{0\} \Rightarrow$ there exists only one optimal solution for $P(\alpha)$.

   $\Rightarrow$ Direction 2: Assume $P(\alpha)$ has a unique optimal solution x, and $Null(A) \cap Null(e^T) \neq \{0\}$. Then, there exists some vector $0 \neq y \in Null(A) \cap Null(e^T)$. We have:

   $$\begin{aligned} & e^T y = 0, e^T x = \alpha \\ & e^T(x-y) = \alpha \\ & ||A(x-y) - b||_2^2 = ||Ax - \underbrace{Ay}_{=0} - b||_2^2 = ||Ax-b||_2^2 = P(\alpha) \end{aligned}$$

   and therefore $(x-y)$ is another optimal solution, in contradiction to the assumption that x is unique. Therefore, unique optimal solution $\Rightarrow Null(A) \cap Null(e^T) = \{0\}$.

2. Recall the KKT conditions:

   $$\begin{aligned} & \nabla f(x) + \mu \nabla h(x) = \nabla (Ax-b)^T(Ax-b) + \mu \nabla (e^T x - \alpha) = 2A^T Ax - 2A^T b + \mu e = 0 \\ & e^T x - \alpha = 0 \end{aligned}$$

The solution to the above is:

$$x^* = (2A^TA)^{-1}(2A^Tb + \mu e)$$

and since A is full rank, $(A^TA)^{-1}$ is well defined. Additionally, due to the constraint:

$$e^Tx^* - \alpha = 0 \Rightarrow e^T(2A^TA)^{-1}(2A^Tb + \mu e) = \alpha$$
$$\Rightarrow \mu e^Te = n\mu = \alpha - e^T(A^TA)^{-1}(A^Tb) \Rightarrow \mu = \frac{1}{n}(\alpha - e^T(A^TA)^{-1}(A^Tb))$$

substituting μ in x^* and we attain that:

$$x^* = \left(2A^TA\right)^{-1}\left(2A^Tb + \frac{1}{n}\left(\alpha - e^T\left(A^TA\right)^{-1}\left(A^Tb\right)\right)e\right)$$

3. Denote x_1 and x_2 as optimal solutions to $P(\alpha_1)$ and $P(\alpha_2)$ respectively, for any α_1, α_2. Then they must obey the constraints of their respective problems:

$$e^Tx_1 = \alpha_1$$
$$e^Tx_2 = \alpha_2$$

and for any $\lambda \in [0, 1]$:

$$\lambda\alpha_1 + (1-\lambda)\alpha_2 = \lambda(e^Tx_1) + (1-\lambda)(e^Tx_2) = e^T(\lambda x_1 + (1-\lambda)x_2)$$

thus the point $\lambda x_1 + (1-\lambda)x_2$ satisfies the constraints of $P(\lambda\alpha_1 + (1-\lambda)\alpha_2)$.
Finally, due to the convexity of $g(x) \triangleq ||Ax-b||_2^2$ and the definition of f as the optimal value, we have:

$$f(\lambda\alpha_1 + (1-\lambda)\alpha_2) \leq g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2) = \lambda f(\alpha_1) + (1-\lambda)f(\alpha_2)$$

which shows that $f(\alpha)$ is convex by definition.

4. First, we show $A^TA + ee^T$ is invertible. For any $x \in \mathbb{R}^n$, we have:

$$x^T(A^TA + ee^T)x = x^TA^TAx + x^Tee^Tx = (Ax)^TAx + (e^Tx)^T(e^Tx) = ||Ax||_2^2 + (e^Tx)^2 \geq 0$$

Additionally, we have:

$$||Ax||_2^2 = 0 \iff Ax = 0 \iff x \in Null(A)$$
$$(e^Tx)^2 = 0 \iff e^Tx = 0 \iff x \in Null(e^T)$$

and since $Null(A) \cap Null(e^T) = \{0\}$, we have that $x^T(A^TA + ee^T)x > 0$ for any $x \in \mathbb{R}^n \backslash \{0\}$ which implies $A^TA + ee^T$ is PD and therefore invertible.

Now, using KKT conditions, x^* is an optimal solution of $P(\alpha)$ if and only if there exists some $\mu \in \mathbb{R}$ such that:

$$\nabla f(x^*) + \mu\nabla h(x^*) = \nabla(Ax^* - b)^T(Ax^* - b) + \mu\nabla(e^Tx^* - \alpha) = 2A^T(Ax^* - b) + \mu e = 0$$
$$e^Tx^* - \alpha = 0$$

adding and subtracting $2ee^Tx^*$ in the first equation and using $e^Tx^* = \alpha$, we have:

$$\begin{aligned}
2A^TAx^* - 2A^Tb + \mu e + 2ee^Tx^* - 2ee^Tx^* &= 0 \\
2(A^TA + ee^T)x^* - 2A^Tb + \mu e - 2\alpha e &= 0 \\
2(A^TA + ee^T)x^* &= 2A^Tb - \mu e + 2\alpha e \\
x^* &= (A^TA + ee^T)^{-1}\left(A^Tb + \left(\alpha - \frac{\mu}{2}\right)e\right)
\end{aligned}$$

Plugging back into the constraint to find μ, we have:

$$\begin{aligned}
e^Tx^* &= \alpha \\
e^T(A^TA + ee^T)^{-1}\left(A^Tb + \left(\alpha - \frac{\mu}{2}\right)e\right) &= \alpha \\
e^T(A^TA + ee^T)^{-1}A^Tb + e^T(A^TA + ee^T)^{-1}\alpha e - e^T(A^TA + ee^T)^{-1}\frac{\mu}{2}e &= \alpha \\
e^T(A^TA + ee^T)^{-1}A^Tb + \alpha e^T(A^TA + ee^T)^{-1}e - \alpha &= \frac{\mu}{2}e^T(A^TA + ee^T)^{-1}e \\
\mu &= 2\frac{e^T(A^TA + ee^T)^{-1}A^Tb}{e^T(A^TA + ee^T)^{-1}e} + 2\alpha
\end{aligned}$$

and finally:

$$x^* = (A^TA + ee^T)^{-1}\left(A^Tb - \frac{e^T(A^TA + ee^T)^{-1}A^Tb}{e^T(A^TA + ee^T)^{-1}e}e\right)$$