553.665 Introduction to Convexity, Fall 2022
Section 9

Hongyu Cheng

November 11, 2022

Discussions on **Theorem 2.7.6** from lecture notes.

Theorem 1 (Theorem 2.7.1 in lecture notes)**.** *The following are equivalent.*

1. *There exists a positive definite matrix* $A \in \mathbb{R}^{d\times d}$ *and* $\mathbf{c} \in \mathbb{R}^d$ *such that*

$$E = \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x}-\mathbf{c})^T A^{-1}(\mathbf{x}-\mathbf{c}) \le 1\}.$$

2. *There exist orthonormal vectors* $\mathbf{b}^1, \dots, \mathbf{b}^d \in \mathbb{R}^d$, $\sigma_1, \dots, \sigma_d > 0$ *and* $\mathbf{c} \in \mathbb{R}^d$ *such that*

$$E = \{\mathbf{c} + \lambda_1 \mathbf{b}^1 + \cdots \lambda_d \mathbf{b}^d : \frac{\lambda_1^2}{\sigma_1^2} + \cdots + \frac{\lambda_d^2}{\sigma_d^2} \le 1\}.$$

Proof. By eigendecomposition of A,

$$A = \begin{bmatrix} \mathbf{b}^1 & \cdots & \mathbf{b}^d \end{bmatrix} \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_d^2 \end{bmatrix} \begin{bmatrix} (\mathbf{b}^1)^T \\ \vdots \\ (\mathbf{b}^d)^T \end{bmatrix} = \underbrace{\begin{bmatrix} \sigma_1\mathbf{b}^1 & \cdots & \sigma_d\mathbf{b}^d \end{bmatrix}}_{:=Q} \underbrace{\begin{bmatrix} (\sigma_1\mathbf{b}^1)^T \\ \vdots \\ (\sigma_d\mathbf{b}^d)^T \end{bmatrix}}_{Q^T},$$

where $\mathbf{b}^1, \dots, \mathbf{b}^d$ are orthonormal, $\sigma_1, \dots, \sigma_d > 0$ since A is positive definite. Then we have

$$\begin{aligned}
E &= \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x}-\mathbf{c})^T A^{-1}(\mathbf{x}-\mathbf{c}) \le 1\} \\
&= \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x}-\mathbf{c})^T Q^{-T} Q^{-1}(\mathbf{x}-\mathbf{c}) \le 1\} \\
&= \{\mathbf{x} \in \mathbb{R}^d : \|Q^{-1}(\mathbf{x}-\mathbf{c})\|_2^2 \le 1\} \\
&= \left\{ \mathbf{c} + \mathbf{x} : \left\| \begin{bmatrix} (\frac{1}{\sigma_1}\mathbf{b}^1)^T \\ \vdots \\ (\frac{1}{\sigma_d}\mathbf{b}^d)^T \end{bmatrix} \mathbf{x} \right\|_2^2 \le 1 \right\} \\
&= \left\{ \mathbf{c} + \mathbf{x} : \frac{((\mathbf{b}^1)^T\mathbf{x})^2}{\sigma_1^2} + \cdots + \frac{((\mathbf{b}^d)^T\mathbf{x})^2}{\sigma_d^2} \le 1 \right\} \\
&= \left\{ \mathbf{c} + \lambda_1\mathbf{b}^1 + \cdots \lambda_d\mathbf{b}^d : \frac{\lambda_1^2}{\sigma_1^2} + \cdots + \frac{\lambda_d^2}{\sigma_d^2} \le 1 \right\}.
\end{aligned}$$

To see the reverse direction, just let $A = QQ^T$, where $Q = \begin{bmatrix} \sigma_1\mathbf{b}^1 & \cdots & \sigma_d\mathbf{b}^d \end{bmatrix}$, and then start from the last equality above to the first. □

Definition 1 (Ellipsoid)**.** Let $A \in \mathbb{R}^{d\times d}$ be a positive definite matrix and $\mathbf{c} \in \mathbb{R}^d$. The set

$$E(A, \mathbf{c}) := \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x} - \mathbf{c})^T A^{-1}(\mathbf{x} - \mathbf{c}) \leq 1\}$$

is called an ellipsoid centered at $\mathbf{c}$.

Theorem 2 (Theorem 2.7.3 in lecture notes)**.** *The following are all true.*

1. *Let $A \in \mathbb{R}^{d\times d}$ be a positive definite matrix and $\mathbf{c} \in \mathbb{R}^d$. Then*

$$\mathrm{vol}(E(A, \mathbf{c})) = \sqrt{\det(A)}\ \mathrm{vol}(B(\mathbf{0}, 1)).$$

2. *Let $\mathbf{b}^1, \dots, \mathbf{b}^d \in \mathbb{R}^d$ be orthonormal vectors, $\sigma_1, \dots, \sigma_d > 0$ and $\mathbf{c} \in \mathbb{R}^d$. Then*

$$\mathrm{vol}(E) = \left(\prod_{i=1}^{d} \sigma_i\right) \mathrm{vol}(B(\mathbf{0}, 1)),$$

where $E = \{\mathbf{c} + \lambda_1\mathbf{b}^1 + \cdots \lambda_d\mathbf{b}^d : \frac{\lambda_1^2}{\sigma_1^2} + \cdots + \frac{\lambda_d^2}{\sigma_d^2} \leq 1\}$.

Proof sketch.

1. Notice that $E(A, \mathbf{c}) = \{\mathbf{x} \in \mathbb{R}^d : \sqrt{(\mathbf{x} - \mathbf{c})^T A^{-1}(\mathbf{x} - \mathbf{c})}\} = \{A^{\frac{1}{2}}\mathbf{y} + \mathbf{c} : \|\mathbf{y}\|_2 \leq 1\}$, then

$$\begin{aligned}
\mathrm{vol}(E(A, \mathbf{c})) &= \int_{E(A,\mathbf{c})} 1 \mathrm{d}x_1 \cdots \mathrm{d}x_d \\
&= \int_{B(\mathbf{0},1)} 1 \cdot \left|\frac{\partial(x_1, \dots, x_d)}{\partial(y_1, \dots, y_d)}\right| \mathrm{d}y_1 \cdots \mathrm{d}y_d \\
&= \int_{B(\mathbf{0},1)} 1 \cdot \left|\det(A^{\frac{1}{2}})\right| \mathrm{d}y_1 \cdots \mathrm{d}y_d \\
&= \sqrt{\det(A)} \int_{B(\mathbf{0},1)} 1 \mathrm{d}y_1 \cdots \mathrm{d}y_d \\
&= \sqrt{\det(A)} \mathrm{vol}(B(\mathbf{0}, 1)).
\end{aligned}$$

2. In this case, $A = QQ^T$, where $Q = \begin{bmatrix}\sigma_1\mathbf{b}^1 & \cdots & \sigma_d\mathbf{b}^d\end{bmatrix}$. Notice that $\sqrt{\det(A)} = \sqrt{\det(Q \cdot Q^T)} = \sqrt{\det(Q) \cdot \det(Q^T)} = \sqrt{\det(Q)^2} = |\det(Q)|$ and the fact that

$$\det(Q) = \det(\begin{bmatrix}\sigma_1\mathbf{b}^1 & \cdots & \sigma_d\mathbf{b}^d\end{bmatrix}) = (\sigma_1 \cdots \sigma_d) \det(\begin{bmatrix}\mathbf{b}^1 & \cdots & \mathbf{b}^d\end{bmatrix}) = (\sigma_1 \cdots \sigma_d)(\pm 1).$$

□

Theorem 3 (Theorem 2.7.4 in lecture notes)**.** *Volume of the unit ball $B(\mathbf{0}, 1)$ in $\mathbb{R}^d$.*

$$\mathrm{vol}(B(\mathbf{0}, 1)) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \sim \frac{1}{\sqrt{\pi n}} \left(\frac{2e\pi}{n}\right)^{\frac{n}{2}}.$$

Moreover,

$$\left(\frac{2}{\sqrt{d}}\right)^d \leq \mathrm{vol}(B(\mathbf{0}, 1)) \leq 2^d.$$

Theorem 4 (Theorem 2.7.6 in lecture notes)**.** *Let $d \geq 2$ and let $E \subseteq \mathbb{R}^d$ be an ellipsoid with center $\mathbf{c} \in \mathbb{R}^d$. Let $H \subseteq \mathbb{R}^d$ be a halfspace and let $0 \leq \beta < \frac{1}{d}$ such that H does not contain $\mathbf{c} + \beta(E - \mathbf{c})$. Then, there exists another ellipsoid E' such that $E \cap H \subseteq E'$ and*

$$\mathrm{vol}(E') \leq e^{-\frac{(1-\beta d)^2}{2(d+1)}} \mathrm{vol}(E).$$

Proof. See the main proof in lecture notes. Some notes on choosing σ and σ' :

Assume $\mathbf{c} = (-t, 0, \ldots, 0) \in \mathbb{R}^d,\ t > 0$, and E' is given by

$$E' = \left\{ \mathbf{x} \in \mathbb{R}^d : \frac{(x_1 + t)^2}{\sigma^2} + \frac{x_2^2}{\sigma'^2} + \cdots + \frac{x_d^2}{\sigma'^2} \leq 1 \right\}.$$

Then imposing the conditions $-\mathbf{e}^1 \in \mathrm{bd}(E'),\ \beta\mathbf{e}^1 \pm \sqrt{1-\beta^2}\mathbf{e}^i \in \mathrm{bd}(E'),\ i = 2, \ldots, d$:

$$\begin{cases} \frac{(-1+t)^2}{\sigma^2} = 1 \\ \frac{(\beta+t)^2}{\sigma^2} + \frac{1-\beta^2}{\sigma'^2} = 1 \end{cases} \implies \begin{cases} t = 1 - \sigma \\ \frac{(\beta+t)^2}{(1-t)^2} + \frac{1-\beta^2}{\sigma'^2} = 1 \end{cases} \implies \sigma'^2 = \frac{(1-t)^2(1-\beta)}{1-\beta-2t}.$$

Then consider the 1-d continuous optimization problem:

$$\min_{t \in (0,1)} \left[(1-t) \left(\frac{(1-t)^2(1-\beta)}{1-\beta-2t} \right)^{\frac{d-1}{2}} = (1-t)^d \frac{(1-\beta)^{\frac{d-1}{2}}}{(1-\beta-2t)^{\frac{d-1}{2}}} \right].$$

Take the logarithm of the objective function and let

$$f(t) = d\log(1-t) - \frac{d-1}{2}\log(1-\beta-2t) + \frac{d-1}{2}(1-\beta).$$

Then

$$\begin{aligned} f'(t) = 0 &\iff \frac{-d}{1-t} - \frac{d-1}{2}\frac{-2}{1-\beta-2t} = 0 \\ &\iff \frac{d-1}{1-\beta-2t} = \frac{d}{1-t} \\ &\iff d - 1 - dt + t = d - \beta d - 2dt \\ &\iff dt + t = 1 - \beta d \\ &\iff t = \frac{1-\beta d}{1+d} \in (0,1). \end{aligned}$$

[Note: Need to check $t = \frac{1-\beta d}{1+d}$ is indeed a minimizer of $f(t)$ over $(0, 1)$]

Therefore,

$$\sigma = 1 - t = \frac{(1+\beta)d}{d+1} \text{ and } \sigma'^2 = \frac{(1-t)^2(1-\beta)}{1-\beta-2t} = \frac{(1-\beta^2)d^2}{d^2-1}.$$

□