As promised in the last lecture, we now give the proof for Proposition 3:

**Proof.** Let \( \bar{a} = \bar{a} / \sqrt{\bar{a}^T B \bar{a}} \), so by Lemma 1,

\[
-1 \leq \bar{a}^T (x - y) \leq +1 \quad \text{for } x \in E. \tag{1}
\]

Now suppose \( x \in E_\alpha \); then

\[
(x - y)^T B^{-1} (x - y) \leq 1. \tag{2}
\]

Also since

\[-1 \leq \bar{a}^T (x - y) \leq -\alpha \quad \text{by equation (1) and the definition of } E_\alpha, \tag{3}\]

we have

\[
(\bar{a}^T (x - y) + \alpha) (\bar{a}^T (x - y) + 1) \leq 0, \quad \text{or} \quad (x - y)^T \bar{a} \bar{a}^T (x - y) + (1 + \alpha) \bar{a}^T (x - y) \leq -\alpha. \tag{4}
\]

From (2) \(\times (1 - \sigma) + (4) \times \sigma \), we get, for any \( 0 \leq \sigma \leq 1 \),

\[
(x - y)^T \left( (1 - \sigma) B^{-1} + \sigma \bar{a} \bar{a}^T \right) (x - y) + (1 + \alpha) \sigma \bar{a}^T (x - y) \leq 1 - \sigma - \sigma \alpha,
\]

\[
\Rightarrow (x - y)^T ((1 - \sigma) B^{-1} + \sigma \bar{a} \bar{a}^T) (x - y) + \frac{(1 + \alpha) \sigma}{2} B \bar{a}) (x - y) + \frac{(1 + \alpha) \sigma}{2} B \bar{a}) \leq 1 - \sigma - \sigma \alpha + \frac{(1 + \alpha)^2 \sigma^2}{4}.
\]

If we set \( y_+ := y - \frac{(1 + \alpha) \sigma}{2} B \bar{a} \) and

\[
B_+^{-1} = \frac{1}{1 - \sigma - \sigma \alpha + \frac{(1 + \alpha)^2 \sigma^2}{4}((1 - \sigma) B^{-1} + \sigma \bar{a} \bar{a}^T)}
\]

\[
= \frac{1 - \sigma}{1 - \sigma - \sigma \alpha + \frac{(1 + \alpha)^2 \sigma^2}{4}} \left( B^{-1} + \frac{\sigma}{1 - \sigma} \bar{a} \bar{a}^T \right)
\]

\[
= \frac{1 - \sigma}{1 - \sigma - \sigma \alpha + \frac{(1 + \alpha)^2 \sigma^2}{4}} \left( B - \frac{\sigma}{1 - \sigma} \bar{a} \bar{a}^T B \right)^{-1} \quad \text{by the Sherman-Morrison-Woodbury formula,}
\]

or

\[
B_+ = \frac{1 - \sigma - \sigma \alpha + \frac{(1 + \alpha)^2 \sigma^2}{4}}{1 - \sigma} (B - \sigma \bar{a} \bar{a}^T B),
\]

this is \((x - y_+)^T B_+^{-1} (x - y_+) \leq 1\). Now plug in

\[
\sigma = \frac{2(1 + n \alpha)}{(1 + n)(1 + \alpha)} \geq 0
\]

with

\[1 - \sigma = \frac{n - 1}{n + 1} \cdot \frac{1 - \alpha}{1 + \alpha} \geq 0.\]
Then
\[ \frac{(1 + \alpha)\sigma}{2} = \frac{1 + n\alpha}{1 + n} = \tau, \]
and after some algebra,
\[ \frac{1 - \sigma - \sigma\alpha + \frac{(1 + \alpha)^2\sigma^2}{4}}{1 - \sigma} = \frac{(1 - \alpha)^2n^2}{n^2 - 1} = \delta. \]
So \( B_+ = \delta(B - \sigma B\bar{a}_a^T B) \) and \( y_+ = y - \tau B\bar{a} \) as in the statement of the proposition. Hence \( E_\alpha \subseteq E_+ \). Also, its volume is
\[
\text{vol}(E_+) = \sqrt{\det B_+} \cdot \text{vol} \text{(unit ball)} \\
= \sqrt{\delta^n \cdot \det B \cdot (1 - \sigma\bar{a}^T B B^{-1} B\bar{a})} \cdot \text{vol} \text{(unit ball)} \quad \text{(by Lemma 2)} \\
= \text{vol}(E) \left[ \frac{n^2}{n^2 - 1} \right]^\frac{n-1}{2} \left( \frac{n - 1}{n + 1} \cdot \frac{1 - \alpha}{1 + \alpha} \right)^{\frac{n}{2}} \\
= \text{vol}(E) \left( \frac{n^2}{n^2 - 1} \right)^\frac{n-1}{2} \left( 1 - \alpha^2 \right)^\frac{n-1}{2} \frac{n}{n + 1} (1 - \alpha). \]
If \( \alpha \geq 0 \), then
\[
\frac{\text{vol}(E_+)}{\text{vol}(E)} \leq \left( 1 + \frac{1}{n^2 - 1} \right)^\frac{n-1}{2} \left( 1 - \frac{1}{n + 1} \right) \\
\leq \left[ \exp \left( \frac{1}{n^2 - 1} \right) \right]^\frac{n-1}{2} \exp \left( -\frac{1}{n + 1} \right) \\
= \exp \left( \frac{1}{2(n + 1)} \right) \exp \left( -\frac{1}{n + 1} \right) \\
= \exp \left( -\frac{1}{2(n + 1)} \right). \]

Here is a sketch of the proof that this is the minimum-volume ellipsoid, in the case \( y = 0, B = I, a = -e_1 \).

Suppose we consider an arbitrary ellipsoid \( \hat{E} := \{ x : \| M x - r \| \leq 1 \} \) with volume \( \frac{1}{\det M} \cdot \text{vol} \text{(unit ball)} \). Choose \( \beta = \sqrt{1 - \alpha^2} \), and consider the points
\[
\alpha e_1 \pm \beta e_j, j = 2, \ldots, n
\]
and $e_1$, all in $E_\alpha$. So, if the columns of $M$ are $m_1, \ldots, m_n$, $||m_1 - r|| \leq 1$ and $||\pm \beta m_j + \alpha m_1 - r|| \leq 1$. So $||\alpha m_1 - r|| =: \gamma \leq 1$, and then we can bound $||m_1||$ and each $||m_j||$ in terms of $\gamma$. But $\det M \leq ||m_1|| \cdot ||m_2|| \cdots ||m_n||$, so we get an upper bound on $\det M$; optimize over $\gamma$ to get a universal bound, which shows $E_+ \in \text{has the minimum volume.}$

**Theorem 1.** If the ellipsoid method is applied to $(f,G)$ where $G = \emptyset$ or $\text{vol}(G) \geq \delta^n$, then if $z_k = \ast$ after $2n(n + 1) \ln \frac{2\sqrt{n}}{\delta}$ steps, $G = \emptyset$, and otherwise, we get $z_k$ with $\epsilon(z_k, f, G) \leq \epsilon$ in $2n(n + 1) \ln \frac{2\sqrt{n}}{\delta}$ steps.

**Proof.** We know each $(E_k, z_k)$ is a localizer. Also, $E_0 = B(nI, 0) = \{x : ||x|| \leq \sqrt{n}\}$ with $\text{vol}(E_0) \leq (2\sqrt{n})^n$. By Proposition 3, every $2(n + 1)$ steps, the volume of $E_k$ is cut by $e$.

To get from volume $(2\sqrt{n})^n$ to $\delta^n$, then, takes

$$2n(n + 1) \ln \left(\frac{2\sqrt{n}}{\delta}\right)$$

steps.

Similarly, we get the volume smaller than $(\delta \epsilon)^n$ within $2n(n + 1) \ln \left(\frac{2\sqrt{n}}{\delta \epsilon}\right)$ steps. \qed

**Comments**

- If $G = C = [-1, 1]^n$, then we can get an $\epsilon$-approximation solution in $2n(n + 1) \ln \left(\frac{1}{\epsilon}\right)$ steps. Exercise (use the fact that $E_0$ is the minimum-volume ellipsoid containing $C$).

- The ellipsoid method is much more general: it shows that “separation $\equiv$ optimization.” We will return to this.

- Forgetting about the details of the scalars, then at each step, the algorithm moves in the direction $-B_k a_k$ (if feasible, $a_k = g(x_k)$). This looks like
  
  - a steepest-descent step ($B_k = I$); or more like
  - a Newton step ($B_k = [\nabla^2 f(x_k)]^{-1}$); or even more like
  - a quasi-Newton step ($B_k \approx [\nabla^2 f(x_k)]^{-1}$, update at each iteration) with a rank-one update.

  This was the viewpoint of N. Shor.

- Proposition 3 can be used to show that every convex body (compact, non-empty interior) in $\mathbb{R}^n$ can be “$n$-rounded”. There exist $B, y$ such that
  
  $$E(n^{-2}B, y) \subseteq C \subseteq E(B, y).$$

  Note that the left-hand side is a copy of the right-hand side, shrunk by a factor of $n$ around its center. This ratio is best possible: let $C$ be a simplex in $\mathbb{R}^n$. 

3