- In the last lecture, we studied the structure of an optimal solution $\hat{z}$ to the problem

$$\hat{z} \in \arg\max_{z \in T^n} z^H C z \qquad (P)$$

In particular, $\hat{z}$ satisfies both the first- and second-order optimality conditions. In this lecture, we show how the structural information on $\hat{z}$ and the definition of GPM can be used to establish the convergence rate of the GPM. Specifically, our goal is to prove the following:

<u>Theorem</u>: Suppose that $\|\Delta\| \le \frac{n^{3/4}}{312}$, $\|\Delta z^*\|_\infty \le \frac{n}{24}$, $\alpha \in [4, \frac{n}{\|\Delta\|})$, and the GPM is initialized by $z^0 = v_c$. Then, there exist $a > 0$ and $\lambda \in (0,1)$ such that for any optimal solution $\hat{z}$ to (P),

$$f(\hat{z}) - f(z^k) \le (f(\hat{z}) - f(z^0))\, \lambda^k,$$

$$d_2(z^k, \hat{z}) \le a \cdot (f(\hat{z}) - f(z^0))^{1/2} \cdot \lambda^{k/2}.$$

Here, $f(z) = z^H C z$.

- The proof of the theorem consists of two main parts. The first is to establish an error bound for (P). Recall that any $z \in T^n$ that satisfies the first- and second-order optimality conditions also satisfies $(\mathrm{Diag}(|\tilde{C}z|) - \tilde{C})z = 0$, where $\tilde{C} = C + \frac{n}{\alpha} I$. This motivates the following formulation of the error bound:

<u>Proposition 1</u>: Let $\Sigma(z) = \mathrm{Diag}(|\tilde{C}z|) - \tilde{C}$, $\rho(z) = \|\Sigma(z) z\|_2$. Under the assumptions of the theorem, for any $z \in T^n$ satisfying $d_2(z, z^*) \le \sqrt{n}/2$ and any optimal solution $\hat{z}$ to (P), we have

$$d_2(z, \hat{z}) \le \frac{8}{n} \rho(z).$$

Before we proceed, several remarks are in order.

(1) Recall that $\hat{z}$ satisfies $\|\hat{z}\|_2^2 = n$ and $(z^*)^H C (z^*) \le \hat{z}^H C \hat{z}$, and hence $d_2(\hat{z}, z^*) \le \frac{4\|\Delta\|}{\sqrt{n}}$. In particular, for $\|\Delta\| \le n/8$, we have $d_2(\hat{z}, z^*) \le \frac{\sqrt{n}}{2}$. Together with Proposition 1, we see that (P) has a unique optimal solution (up to a global phase).

(2) We have previously shown that $d_2(z^k, z^*) \le \sqrt{n}/2$ whenever $\|\Delta\| \le n/16$; see the proof of Theorem 1 in the notes of Week 12. This shows that the error bound in Proposition 1 applies to the entire sequence of iterates $\{z^k\}_{k\ge 0}$.

(3) Recall that if $z \in \mathbb{T}^n$ satisfies the first- and second-order optimality conditions, then $\rho(z) = 0$. Thus, all such points that also satisfy $d_2(z, z^*) \le \sqrt{n}/2$ are optimal for (P). In other words, the first- and second-order optimality conditions are sufficient for optimality.

The second part of the proof of the theorem is to establish the following properties of the GPM:

<u>Proposition 2</u>: Under the assumptions of the theorem, there exist $a_0, a_1, a_2 > 0$ such that for any optimal solution $\hat{z}$ to (P), the following hold:

(a) (Sufficient ascent) $f(z^{k+1}) - f(z^k) \ge a_0 \cdot \|z^{k+1} - z^k\|_2^2$

(b) (cost-to-go estimate) $f(\hat{z}) - f(z^k) \le a_1 \cdot d_2(z^k, \hat{z})^2$

(c) (safeguard) $\rho(z^k) \le a_2 \cdot \|z^{k+1} - z^k\|_2$.

- Armed with Propositions 1 and 2, we are ready to prove the theorem:

Proof of the Theorem:

$$
\begin{aligned}
f(\hat{z}) - f(z^{k+1}) &= \left(f(\hat{z}) - f(z^k) \right) - \left(f(z^{k+1}) - f(z^k) \right) \\
&\le a_1 \cdot d_2(z^k, \hat{z})^2 - \left(f(z^{k+1}) - f(z^k) \right) && \text{(cost-to-go)} \\
&\le \frac{64 a_1}{n^2} \rho(z^k)^2 - \left(f(z^{k+1}) - f(z^k) \right) && \text{(error bound)} \\
&\le \frac{64 a_1 a_2^2}{n^2} \| z^{k+1} - z^k \|_2^2 - \left(f(z^{k+1}) - f(z^k) \right) && \text{(safeguard)} \\
&\le \left(\frac{64 a_1 a_2^2}{a_0 n^2} - 1 \right) \left(f(z^{k+1}) - f(\hat{z}) + f(\hat{z}) - f(z^k) \right). && \text{(sufficient ascent)}
\end{aligned}
$$

Since $f(\hat{z}) \geq f(z^k)$ for all k, we may assume without loss that $a' = \dfrac{64 a_1 a_2^2}{a_0 n^2} > 1$. Hence,

$$
f(\hat{z}) - f(z^{k+1}) \le \frac{a'-1}{a'} \left(f(\hat{z}) - f(z^k) \right),
$$

which yields $f(\hat{z}) - f(z^{k+1}) \le \left(f(\hat{z}) - f(z^0) \right) \lambda^k$. Furthermore,

$$
\begin{aligned}
d_2(z^k, \hat{z})^2 &\le \frac{64}{n^2} \rho(z^k)^2 && \text{(error bound)} \\
&\le \frac{64 a_2^2}{n^2} \| z^{k+1} - z^k \|_2^2 && \text{(safeguard)} \\
&\le \frac{64 a_2^2}{a_0 n^2} \left(f(z^{k+1}) - f(z^k) \right) && \text{(sufficient ascent)} \\
&\le \frac{64 a_2^2}{a_0 n^2} \left(f(\hat{z}) - f(z^k) \right) \\
&\le \frac{64 a_2^2}{a_0 n^2} \left(f(\hat{z}) - f(z^0) \right) \lambda^k .
\end{aligned}
$$

This completes the proof.

- The above result once again demonstrates the power of the error bound-based convergence rate analysis framework.

- Let us now proceed to the proof of Proposition 1.

Proof of Proposition 1

Observe that

$$\rho(z) = \|\Sigma(z) z\|_2 \geq \|\Sigma(\hat{z}) z\|_2 - \|(\Sigma(z) - \Sigma(\hat{z})) z\|_2 .$$

Let $\hat{\theta} = \underset{\theta \in [0,2\pi)}{\arg\min} \|z - e^{i\theta}\hat{z}\|_2$, $\hat{\theta}^* = \underset{\theta \in [0,2\pi)}{\arg\min} \|\hat{z} - e^{i\theta} z^*\|_2$

and note that $\tilde{C} = C + \frac{n}{\alpha} I = z^*(z^*)^H + \Delta + \frac{n}{\alpha} I$. We first bound

$$\|(\Sigma(z) - \Sigma(\hat{z})) z\|_2 = \|(\mathrm{Diag}(|\tilde{C} z|) - \mathrm{Diag}(|\tilde{C}\hat{z}|)) z\|_2$$

$$= \left(\sum_{j=1}^{n} \left| \left(|(\tilde{C} z)_j| - |(\tilde{C}\hat{z})_j| \right) z_j \right|^2 \right)^{1/2}$$

$$= \| \, |\tilde{C} e^{-i\hat{\theta}} z| - |\tilde{C}\hat{z}| \, \|_2 \qquad (\text{since } |z_j| = 1)$$

$$\leq \|\tilde{C}(e^{-i\hat{\theta}} z - \hat{z})\|_2$$

$$\leq \|z^*(z^*)^H (e^{-i\hat{\theta}} z - \hat{z})\|_2 + \|\Delta(e^{-i\hat{\theta}} z - \hat{z})\|_2 + \frac{n}{\alpha} \|e^{-i\hat{\theta}} z - \hat{z}\|_2$$

$$\leq \sqrt{n} \cdot |(z^*)^H (e^{-i\hat{\theta}} z - \hat{z})\|_2 + \left(\|\Delta\| + \frac{n}{\alpha} \right) d_2(z, \hat{z}). \qquad -(+)$$

By definition of $\hat{\theta}$, we have $\hat{z}^H(e^{-i\hat{\theta}} z) = |\hat{z}^H z|$. Thus,

$$d_2(z,\hat{z})^2 = \|e^{-i\hat{\theta}} z - \hat{z}\|_2^2 = 2(n - |\hat{z}^H z|). \qquad -(*)$$

Recalling that $d_2(\hat{z}, z^*) \leq \frac{4\|\Delta\|}{\sqrt{n}}$, we obtain

$$|(z^*)^H (e^{-i\hat{\theta}} z - \hat{z})| \leq |(z^* - e^{-i\hat{\theta}^*}\hat{z})^H (e^{-i\hat{\theta}} z - \hat{z})| + |(e^{-i\hat{\theta}^*}\hat{z})^H (e^{-i\hat{\theta}} z - \hat{z})|$$

$$\leq \|z^* - e^{-i\hat{\theta}^*}\hat{z}\|_2 \cdot \|e^{-i\hat{\theta}} z - \hat{z}\|_2 + \left| |\hat{z}^H z| - n \right|$$

$$\leq \frac{4\|\Delta\|}{\sqrt{n}} \cdot d_2(z,\hat{z}) + \frac{1}{2} d_2(z,\hat{z})^2 \qquad -(++)$$

It follows from (+) and (++) that

$$\|(\Sigma(z) - \Sigma(\hat{z})) z\|_2 \leq \left(5\|\Delta\| + \frac{n}{\alpha}\right) d_2(z,\hat{z}) + \frac{\sqrt{n}}{2} d_2(z,\hat{z})^2 .$$

Next, we show that $\Sigma(\hat{z}) \succeq 0$ and $\Sigma(\hat{z})$ is pd on the orthogonal complement of $\mathrm{span}(\hat{z})$. Indeed, we have $\Sigma(\hat{z})\hat{z} = 0$ because $\hat{z}$ satisfies both the first- and second-order optimality conditions.

Next, let $\hat{u} = (I - \frac{1}{n}\hat{z}\hat{z}^H)(e^{-i\hat{\theta}}z - \hat{z})$ be the projection of $e^{-i\hat{\theta}}z - \hat{z}$ onto the orthogonal complement of span $(\hat{z})$. Then, we have $\hat{u}^H\hat{z} = 0$ and

$$\|\hat{u}\|_2 \geq \|e^{-i\hat{\theta}}z - \hat{z}\|_2 - \frac{1}{n}\|\hat{z}\hat{z}^H(e^{-i\hat{\theta}}z - \hat{z})\|_2$$

$$= d_2(z,\hat{z}) - \frac{1}{2\sqrt{n}}d_2(z,\hat{z})^2 \quad \text{(by (*) and } \hat{z}^H(e^{-i\hat{\theta}}z) = |\hat{z}^H z|). \quad -(\Delta)$$

Also, by definition of $\hat{\theta}^*$, we have

$$(z^*)^H(e^{-i\hat{\theta}^*}\hat{z}) = |(z^*)^H\hat{z}| \quad \text{and} \quad |(z^*)^H\hat{z}| = n - \frac{1}{2}\|\hat{z} - e^{i\hat{\theta}^*}z^*\|_2^2. \quad -(**)$$

It follows that

$$\hat{u}^H\Sigma(\hat{z})\hat{u} = \hat{u}^H(\text{Diag}(|\tilde{C}\hat{z}|) - \tilde{C})\hat{u} = \hat{u}^H(\text{Diag}(|C\hat{z}|) - C)\hat{u}$$

(since $\hat{z}$ satisfies $|(\tilde{C}\hat{z})_j| = (\tilde{C}\hat{z})_j\bar{\hat{z}}_j = (C\hat{z})_j\bar{\hat{z}}_j + \frac{n}{\alpha} = |(C\hat{z})_j| + \frac{n}{\alpha}$; see pp. 10–11 of the notes of Week 10)

$$\geq \sum_{j=1}^n |(C\hat{z})_j|\cdot|\hat{u}_j|^2 - |(z^*)^H\hat{u}|^2 - \hat{u}^H\Delta\hat{u} \quad (\text{since } C = z^*(z^*)^H + \Delta)$$

$$\geq (|(z^*)^H\hat{z}| - \|\Delta\hat{z}\|_\infty)\cdot\|\hat{u}\|_2^2 - |\hat{u}^H(z^* - e^{-i\hat{\theta}^*}\hat{z})|^2 - \|\Delta\|\cdot\|\hat{u}\|_2^2$$

$$\geq \left(n - \|\Delta\hat{z}\|_\infty - \|\Delta\| - \frac{3}{2}\|z^* - e^{-i\hat{\theta}^*}\hat{z}\|_2^2\right)\cdot\|\hat{u}\|_2^2 \quad (\text{by } (**))$$

$$\geq \left(n - \|\Delta\hat{z}\|_\infty - \|\Delta\| - \frac{24\|\Delta\|^2}{n}\right)\cdot\|\hat{u}\|_2^2 \quad \left(\text{since } d_2(\hat{z},z^*) \leq \frac{4\|\Delta\|}{\sqrt{n}}\right).$$

In particular, since $\Sigma(\hat{z})\hat{z} = 0$,

$$\|\Sigma(\hat{z})z\|_2 = \|\Sigma(\hat{z})(e^{-i\hat{\theta}}z - \hat{z})\|_2 = \|\Sigma(\hat{z})\hat{u}\|_2$$

$$\geq \left(n - \|\Delta\hat{z}\|_\infty - \|\Delta\| - \frac{24\|\Delta\|^2}{n}\right)\left(d_2(z,\hat{z}) - \frac{1}{2\sqrt{n}}d_2(z,\hat{z})^2\right) \quad (\text{by } (\Delta))$$

Since $d_2(z,z^*) \leq \frac{\sqrt{n}}{2}$ by assumption, we have

$$d_2(z,\hat{z}) \leq d_2(z,z^*) + d_2(\hat{z},z^*) \leq \frac{\sqrt{n}}{2} + \frac{4\|\Delta\|}{\sqrt{n}},$$

which implies

$$d_2(z,\hat{z})^2 \leq \left(\frac{\sqrt{n}}{2} + \frac{4\|\Delta\|}{\sqrt{n}}\right)d_2(z,\hat{z}).$$

Moreover,

$$\begin{aligned}\|\Delta\hat{z}\|_\infty &\le \|\Delta z^*\|_\infty + \|\Delta(e^{-i\hat{\theta}^*}\hat{z} - z^*)\|_\infty \\ &\le \|\Delta z^*\|_\infty + \|\Delta\| \cdot d_2(\hat{z}, z^*) \\ &\le \|\Delta z^*\|_\infty + \frac{4\|\Delta\|^2}{\sqrt{n}}.\end{aligned}$$

Putting everything together and invoking the assumptions, we have

$$\rho(z) \geq \|\Sigma(\hat{z})z\|_2 - \|(\Sigma(z) - \Sigma(\hat{z}))z\|_2 \geq \frac{n}{8} d_2(z, \hat{z}),$$

as desired.

- Next, let us turn to the proof of Proposition 2.

<u>Proof of Proposition 2</u>

(a) Recall that $\tilde{C} = C + \frac{n}{\alpha}I$. We compute

$$f(z^{k+1}) - f(z^k) = (z^{k+1} - z^k)^H \tilde{C}(z^{k+1} - z^k) - 2(z^k)^H \tilde{C} z^k + 2\,\mathrm{Re}[(z^{k+1})^H \tilde{C} z^k].$$

By definition of z^{k+1}, we have $(z^{k+1})^H \tilde{C} z^k = \left(\frac{\tilde{C}z^k}{|\tilde{C}z^k|}\right)^H \tilde{C} z^k$, which, as one would observe, is a real number. Moreover,

$$z^{k+1} \in \underset{z \in \mathbb{T}^n}{\mathrm{argmax}}\ \mathrm{Re}(z^H \tilde{C} z^k),$$

which implies that $\mathrm{Re}[(z^{k+1})^H \tilde{C} z^k] \geq (z^k)^H \tilde{C} z^k$. Hence,

$$f(z^{k+1}) - f(z^k) \geq (z^{k+1} - z^k)^H \tilde{C}(z^{k+1} - z^k) \geq a_0 \cdot \|z^{k+1} - z^k\|_2^2$$

with $a_0 = \lambda_{min}(\Delta + \frac{n}{\alpha}I) > 0$ by assumption on α.

(b) Let $\hat{\theta}_k = \underset{\theta \in [0, 2\pi)}{\mathrm{argmin}} \|z^k - e^{i\theta}\hat{z}\|_2$. Then,

$$\begin{aligned}f(\hat{z}) - f(z^k) &= \hat{z}^H \tilde{C} \hat{z} - (z^k)^H \tilde{C} z^k \\ &= \|\tilde{C}\hat{z}\|_1 - (z^k)^H \tilde{C} z^k\end{aligned}$$

(by Proposition 5 of Week 10's notes)

$$= (z^k)^H [\mathrm{Diag}(|\tilde{C}\hat{z}|) - \tilde{C}] z^k$$

(since $(z^k)^H \mathrm{Diag}(|\tilde{C}\hat{z}|) z^k = \sum_i |(\tilde{C}\hat{z})_i| = \|\tilde{C}\hat{z}\|_1$)

$$= (e^{-i\hat{\theta}_k} z^k - \hat{z})^H \left(\mathrm{Diag}(|\tilde{C}\hat{z}|) - \tilde{C} \right) (e^{-i\hat{\theta}_k} z^k - \hat{z})$$

$$\left(\text{since } \left[\mathrm{Diag}(|\tilde{C}\hat{z}|) - \tilde{C}\right]\hat{z} = 0\right)$$

$$\leq \left(\|\tilde{C}\| + \|\tilde{C}\hat{z}\|_\infty \right) \cdot d_2(z^k, \hat{z})^2.$$

We bound

$$\|\tilde{C}\| + \|\tilde{C}\hat{z}\|_\infty \leq \|C\| + \|C\hat{z}\|_\infty + \frac{2n}{\alpha}$$

$$\leq \|z^*(z^*)^H\| + \|\Delta\| + \|z^*(z^*)^H \hat{z}\|_\infty + \|\Delta \hat{z}\|_\infty + \frac{2n}{\alpha}.$$

Observe that $\|z^*(z^*)^H\| = n$, $\|z^*(z^*)^H \hat{z}\|_\infty = |(z^*)^H \hat{z}| \cdot \|z^*\|_\infty \leq n$

and

$$\|\Delta\hat{z}\|_\infty \leq \|\Delta z^*\|_\infty + \|\Delta(e^{-i\hat{\theta}^*}\hat{z} - z^*)\|_\infty$$

$$\leq \|\Delta z^*\|_\infty + \|\Delta\| \cdot d_2(\hat{z}, z^*) \leq \|\Delta z^*\|_\infty + \frac{4\|\Delta\|^2}{\sqrt{n}}.$$

Thus,

$$\|\tilde{C}\| + \|\tilde{C}\hat{z}\|_\infty \leq 2n + \|\Delta\| + \|\Delta z^*\|_\infty + \frac{4\|\Delta\|^2}{\sqrt{n}} + \frac{2n}{\alpha} < 3n,$$

where the last inequality is due to our assumptions. Hence, we have

$$f(\hat{z}) - f(z^k) \leq a_1 \cdot d_2(z^k, \hat{z})^2$$

for some $a_1 \in (0, 3n)$.

(c) By definition of z^{k+1}, we have

$$\mathrm{Diag}(|\tilde{C}z^k|)(z^{k+1} - z^k) = \left[\tilde{C} - \mathrm{Diag}(|\tilde{C}z^k|)\right] z^k.$$

Hence,

$$\rho(z^k) = \|\mathrm{Diag}(|\tilde{C}z^k|)(z^{k+1} - z^k)\|_2 \leq \|\tilde{C}z^k\|_\infty \cdot \|z^{k+1} - z^k\|_2.$$

Recall that $d_2(z^k, z^*) \leq \frac{\sqrt{n}}{2}$ (See Remark 2 on p. 2). Let $\theta_k^* = \underset{\theta\in[0,2\pi)}{\mathrm{argmin}} \|z^k - e^{i\theta} z^*\|_2$. Then, by our assumptions,

$$\|\tilde{C}z^k\|_\infty \leq \|z^*(z^*)^H z^k\|_\infty + \|\Delta z^k\|_\infty + \frac{n}{\alpha}$$

$$\leq |(z^*)^H z^k| + \|\Delta(e^{-i\theta_k^*} z^k - z^*)\|_\infty + \|\Delta z^*\|_\infty + \frac{n}{\alpha}$$

$$\leq n + \|\Delta\| \cdot d_2(z^k, z^*) + \|\Delta z^*\|_\infty + \frac{n}{\alpha}$$

$$< \frac{3}{2} n^{5/4}.$$

Hence, $\rho(z^k) \leq a_2 \cdot \|z^{k+1} - z^k\|_2$ for some $a_2 \in (0, \frac{3}{2} n^{5/4})$.