LECTURE 7

- PROBLEM SETTING: MATRIX SENSING IN THE NOISELESS SETTING.

$$y = \mathcal{A}(X^*), \quad \text{WHERE } X^* \text{ IS RANK-}r. \quad X^* \succcurlyeq 0.$$

REST CONFIGURATION:

i) $\mathcal{A}(\cdot): \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^m$, $m \ll n^2$ AND $\mathcal{A}(\cdot)$ SATISFIES RIP.

ii) $y \in \mathbb{R}^m$ SET OF MEASUREMENTS.

WE ARE INTERESTED IN THE BEHAVIOR OF THE NON-CONVEX OBJECTIVE:

$$\min_{U \in \mathbb{R}^{n\times r}} \left\{ f(U) := \| y - \mathcal{A}(UU^T) \|_2^2 \right\}$$

THE ALGORITHM TO USE IS: $U_{t+1} = U_t - \eta \nabla f(U_t U_t^T) \cdot U_t$

- QUESTIONS:
  1. CHARACTERIZE FIRST-ORDER STATIONARY POINT.
  2. FIGURE OUT WHETHER ANY LOCAL MINIMA ARE DIFFERENT THAN THE GLOBAL OPTIMAL.
  3. FIGURE OUT WHETHER THE REST OF STATIONARY POINTS (= SADDLE POINTS) ARE STRICT.

- BY ASSUMPTION OF RIP:

$$(1-\delta)\cdot \|X\|_F^2 \le \frac{1}{m} \|\mathcal{A}(X)\|_2^2 \le (1+\delta) \|X\|_F^2, \quad \forall X \text{ rank-}r \text{ (OR LESS)}$$

$$\Rightarrow (1-\delta)\|X\|_F^2 \le \frac{1}{m} \sum_{i=1}^m \langle A_i, X\rangle^2 \le (1+\delta)\|X\|_F^2$$

A CONSEQUENCE OF THE RIP:

GIVEN TWO $n\times n$ RANK-r MATRICES X AND Y, AND GIVEN THAT $\mathcal{A}$ SATISFIES RIP, THE FOLLOWING HOLDS:

$$\left| \frac{1}{m} \sum_{i=1}^m \langle A_i, X\rangle \langle A_i, Y\rangle - \langle X, Y\rangle \right| \le \delta \cdot \|X\|_F \cdot \|Y\|_F$$

- STATIONARY POINTS:

$$\nabla f(U) = 0 \Rightarrow -2\mathcal{A}^+\big(y - \mathcal{A}(UU^T)\big)\cdot U = 0$$

$$\Rightarrow \mathcal{A}^+\big(y - \mathcal{A}(UU^T)\big)\cdot U = 0$$

$$\Rightarrow \sum_{i=1}^m \big(y - \mathcal{A}(UU^T)\big)_i \cdot A_i \cdot U = 0$$

$$\Rightarrow \sum_{i=1}^{m} \left(\mathcal{A}(x^*) - \mathcal{A}(UU^T)\right)_i \, A_i \cdot U = 0$$

$$\Rightarrow \sum_{i=1}^{m} \langle A_i, UU^T - U^*U^{*T} \rangle \cdot A_i \cdot U = 0. \longleftarrow$$

THIS IS THE CONDITION THAT ALL STATIONARY POINTS SATISFY.

FOR ALL STATIONARY POINTS, THE FOLLOWING HOLD: LET $U = QR$ BE THE QR-DECOMPOSITION OF U. LET MATRIX ZQR^{-1} FOR SOME Z MATRIX IN $n \times r$. THEN: $\hookrightarrow \|Z\|_F \leq 1$.

$$\frac{1}{m} \sum_{i=1}^{m} \langle A_i, UU^T - U^*U^{*T} \rangle \cdot \langle A_i \cdot U, ZQR^{-1} \rangle = 0. \Rightarrow$$

$$\frac{1}{m} \sum_{i=1}^{m} \langle A_i, UU^T - U^*U^{*T} \rangle \langle A_i, ZQR^{-1}U^T \rangle = 0 \Rightarrow$$

$$\frac{1}{m} \sum_{i=1}^{m} \langle A_i, UU^T - U^*U^{*T} \rangle \cdot \langle A_i, ZQR^{-1}RQ^T \rangle = 0 \Rightarrow$$

$$\frac{1}{m} \sum_{i=1}^{m} \langle A_i, UU^T - U^*U^{*T} \rangle \langle A_i, QQ^T Z^T \rangle = 0.$$

BY RIP:

$$\left| \langle UU^T - U^*U^{*T}, QQ^T Z^T \rangle \right| \leq \delta \, \|UU^T - U^*U^{*T}\|_F \cdot \|QQ^T \cdot Z^T\|_F$$

$\forall Z$ S.T. $\|Z\|_F \leq 1$. THUS IT HOLDS ALSO FOR $\sup_{Z: \|Z\|_F \leq 1} \langle (UU^T - U^*U^{*T}) QQ^T, Z \rangle = \|(UU^T - U^*U^{*T}) QQ^T\|$

THUS: $\|(UU^T - U^*U^{*T}) QQ^T\|_F \leq \delta \cdot \|UU^T - U^*U^{*T}\|_F$

- SECOND-ORDER CONDITIONS

$\nabla^2 f(U) \in \mathbb{R}^{nr \times nr}$. IN ORDER TO CHECK THE SECOND-ORDER DERIVATIVE TEST. WE MIGHT ALSO COMPUTE THE QUADRATIC FORM:

$$\text{vec}(Z)^T \cdot \nabla^2 f(U) \cdot \text{vec}(Z) \geq 0., \quad \text{FOR } Z \in \mathbb{R}^{n \times r}$$

MOREOVER, WE CAN USE THE HESSIAN-VECTOR APPROXIMATION:

$$\nabla^2 f(U) \cdot \text{vec}(Z) = \lim_{t \to 0} \left[\frac{\nabla f(U + tZ) - \nabla f(U)}{t} \right]$$

GIVEN THE ABOVE, FOR ANY MATRIX $Z \in \mathbb{R}^{n \times r}$ WE HAVE:
AND FOR LOCAL MINIMUM U,

$$\text{vec}(Z)^T \nabla^2 f(U) \cdot \text{vec}(Z) =$$

$$= \text{vec}(Z)^T \cdot \lim_{t \to 0} \left[\frac{\nabla f(U + tZ) - \nabla f(U)}{t} \right]$$

(WE COMPUTE $\nabla f(U+tZ), \nabla f(U)$)

$$= \sum_{i=1}^{m} \left[4 \langle A_i, UZ^T \rangle^2 + 2 \langle A_i, UU^T - U^* U^{*T} \rangle \langle A_i, ZZ^T \rangle \right]$$

SET $Z = U - U^* R$. THEN, USING FIRST-ORDER OPTIMALITY CONDITION:

$$= \sum_{i=1}^{m} \left[4 \langle A_i, U \cdot (U - U^* R)^T \rangle^2 - 2 \langle A_i, UU^T - U^* U^{*T} \rangle^2 \right]$$

≥ 0 (AS A LOCAL MINIMUM).

MOREOVER, WE OBTAIN (OUT OF THE SCOPE OF THIS LECTURE):

i) $\|U(U - U^* R)\|_F^2 \geq \frac{1-\delta}{2(1+\delta)} \|UU^T - U^* U^{*T}\|_F^2$

ii) $\|U(U - U^* R)\|_F^2 \leq \frac{1}{8} \|UU^T - U^* U^{*T}\|_F^2 + \frac{34}{8} \|(UU^T - U^* U^{*T}) Q Q^T\|_F^2$

COMBINING ALL THE ABOVE, ASSUME $UU^T \neq U^* U^{*T}$. THEN:

$$\left(\frac{1-\delta}{2(1+\delta)} - \frac{1}{8} \right) \|UU^T - U^* U^{*T}\|_F^2 \leq \frac{34}{8} \delta^2 \|(UU^T - U^* U^{*T})\|_F^2$$

IF $\delta \leq 1/5$, THEN THE INEQUALITY HOLDS ONLY IF $UU^T = U^* U^{*T}$.

THIS PROVES THAT FOR $\mathcal{A}(\cdot)$ WITH RIP CONSTANT $\delta \leq 1/5$,

ALL LOCAL MINIMA = GLOBAL MINIMA

FURTHER ONE CAN SHOW THAT:

$$\lambda_{min} \left(\frac{1}{m} \nabla^2 f(U) \right) \leq -\frac{4}{5} \sigma_r(X^*) \longrightarrow \text{STRICT SADDLE PROPERTY}$$