Lemma 4.1

Given 2 $n \times n$ rank-r matrices X and Y and a $(2r, \delta)$-RIP measurement operator $\mathcal{A}$, $\Rightarrow$

not necessary

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

(pf: S163, Lemma 2.1)

only need linear combination X, Y, rank $\le 2r$

WLOG $\|X\|_F = \|Y\|_F = 1$

parallelogram Law: $4\langle C, D\rangle = \|C+D\|^2 - \|C-D\|^2$

$\Rightarrow 4\langle \mathcal{A}(X), \mathcal{A}(Y)\rangle = \|\mathcal{A}(X+Y)\|^2 - \|\mathcal{A}(X-Y)\|^2$, $\quad 4\langle X, Y\rangle = \|X$

.. see page 3

- 1st order optimality $\quad U \in \mathbb{R}^{n \times r}$

$$\nabla f(U) = 0 \Rightarrow \sum_i \langle A_i, UU^T - U^* U^{*T} \rangle A_i U = 0 \qquad (2)$$

$A_i \in \mathbb{R}^{n \times n}$

$A_i U \in \mathbb{R}^{n \times r}$

Lemma 4.2

$\mathcal{A}$ satisfied $(2r, \delta)$-RIP

then $$\|(UU^T - U^*U^{*T}) QQ^T\|_F \le \delta \|UU^T - U^*U^{*T}\|_F \qquad (3)$$

Q is an orthonormal matrix that span column space of U

[i.e. U is close to U^* in subspace spanned by column of U]

<pf> Let $U = QR$, for some orthonormal Q, $R \in \mathbb{R}^{r \times r}$ is invertiable

consider any matrix of form $Z' = ZQR^{-1T}$ for arbitrary Z ? $\in \mathbb{R}^{n \times n}$

$$S = \sum_{i=1}^{m} \langle A_i, UU^T - X^* \rangle A_i U = 0_{n\times r}$$

$$\Rightarrow \quad \langle \text{LHS}, Z' \rangle = \langle 0_{n\times r}, Z' \rangle = 0$$

$$\Rightarrow \quad \sum_{i=1}^{m} \langle A_i, UU^T - X^* \rangle \langle A_i U, Z' \rangle \qquad \cdots (3.2)$$

$$\langle A_i U, Z' \rangle = tr((A_i U)^T Z') = tr(U^T A_i^T Z')$$

$$= tr(A_i^T Z' U^T) \quad (\text{property of trace})$$

$$= tr(A_i^T (U Z'^T))$$

$$= \langle A_i, U Z'^T \rangle$$

$$U Z'^T = QR \cdot (Z Q R^{-1 T})^T = QR \cdot R^{-1} Q^T Z^T = Q Q^T Z^T$$

Let $Y_1 := UU^T - X^*$, rank $\le$ rank(UU^T) + rank(X^*) $\le \underline{2r}$?

$Y_2 := QQ^T Z^T$ rank $\le r$, since $Z \in \mathbb{R}^{r\times n}$

Column of $Y_1 \in$ span$(U) \cup$ span(U_*)

col $Y_2 \in$ span(U)

Since span(Q) = span(U)

$\Rightarrow$ col $(Y_1 \pm Y_2) \in$ span$(U) \cup S(U_*)$

$\Rightarrow$ rank $\le 2r$

by (1)

$$\Big| \underbrace{\sum_{i=1}^{m} \langle A_i, Y_1 \rangle \langle A_i, Y_2 \rangle}_{=0,\ \text{by } (3.2)} - m \langle Y_1, Y_2 \rangle \Big| \le m\delta \|Y_1\|_F \|Y_2\|_F$$

$$\Rightarrow \quad |-m \langle Y_1, QQ^T Z^T \rangle| \le m\delta \|Y_1\|_F \|QQ^T Z^T\|_F$$

$$\Rightarrow \quad |\langle Y_1, QQ^T Z^T \rangle| \le \delta \|Y_1\|_F \|QQ^T Z^T\|_F$$

$$\Rightarrow \quad \langle Y_1 QQ^T, Z^T \rangle \le \delta \|Y_1\|_F \|QQ^T Z^T\|_F, \quad \forall Z \in \mathbb{R}^{r\times n}$$

$$\Rightarrow \quad \|(UU^T - U_* U_*^T) QQ^T\|_F \le \delta \|Y_1\|_F$$

Pf of Lemma 4.1.

WLOG $\|X\|_F = \|Y\|_F = 1$

parallelogram Inq: $4\langle C, D\rangle = \|C+D\|^2 - \|C-D\|^2$

$\Rightarrow 4\langle \mathcal{A}(X), \mathcal{A}(Y)\rangle = \|\mathcal{A}(X+Y)\|^2 - \|\mathcal{A}(X-Y)\|^2,$

$4\langle X, Y\rangle = \|X+Y\|^2 - \|X-Y\|^2$

$\Rightarrow 4\Delta := 4(\langle \mathcal{A}(X), \mathcal{A}(Y)\rangle - \langle X, Y\rangle) = (\|\mathcal{A}(X+Y)\|^2 - \|X+Y\|^2) - (\|\mathcal{A}(-)\|^2 - \|-\|^2)$

by RIP. $\forall Z$, with $\mathrm{rank}(Z) \le 2r$, we have

$(1-\delta)\|Z\|_F^2 \le \|\mathcal{A}(Z)\|^2 \le (1+\delta)\|Z\|_F^2$

$\Rightarrow \left|\|\mathcal{A}(Z)\|^2 - \|Z\|_F^2\right| \le \delta \|Z\|_F^2$

$\Rightarrow 4|\Delta| \le \delta\|X+Y\|_F^2 + \delta\|X-Y\|_F^2$ triangular Inq

$= \delta \cdot 2(\|X\|_F^2 + \|Y\|_F^2)$

$\le \delta \cdot 4 \cdot \|X\|_F \|Y\|_F$. □

Lemma 4.3 [Hessian computation]

U be 1-order critical pt of $f(U)$

$\forall R \in \mathbb{R}^{r\times r}$, $\Delta_j = \Delta e_j e_j^T$ $(\Delta = U - U^* R)$

$$\sum_{j=1}^{r} vec(\Delta_j)^T [\nabla^2 f(U)]\, vec(\Delta_j) = \sum_{i=1}^{m}\left(\sum_{j=1}^{r} 4\langle A_i, U\Delta_j^T\rangle^2 - 2\langle A_i, UU^T - U_* U_*^T\rangle^2\right)$$

<pf>

$\forall Z$

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

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

Since $\nabla f(U) = 2 \sum_{i=1}^{m} [\langle A_i, UU^T - X^*\rangle] A_i U$

$$\lim \left[\frac{\nabla f(U + t\Delta_j) - \nabla f(U)}{t} \right]$$

$$= 2 \sum_{i=1}^{m} \lim_{t\to 0} \frac{1}{t} \left[\langle A_i, (U + t\Delta_j)(U + t\Delta_j)^T - X^* \rangle A_i (U + t\Delta_j) - \langle A_i, UU^T - X^* \rangle A_i U \right]$$

$$(U + t\Delta_j)(U + t\Delta_j)^T = UU^T + t(U\Delta_j^T + \Delta_j U^T) + t^2 \Delta_j \Delta_j^T$$

$$= 2 \sum_{i=1}^{m} \lim_{t\to 0} \frac{1}{t} \left\{ \left[\langle A_i, UU^T - X^*\rangle + \langle A_i, t(\cdots)\rangle + \langle A_i, t^2 \Delta_j \Delta_j^T\rangle \right] (A_i U + A_i t\Delta_j) - \langle A_i, UU^T - X^*\rangle A_i U \right\}$$

$$2 \sum_{i=1}^{m} \left\{ \langle A_i, (U\Delta_j^T + \Delta_j U^T)\rangle A_i U + \langle A_i, UU^T - X^*\rangle A_i \Delta_j \right\}$$

$\Rightarrow$

$$\mathrm{vec}(\Delta_j)^T [\nabla^2 f(U)] \mathrm{vec}(\Delta_j)$$

$$= 2 \sum_{i=1}^{m} \left(\langle A_i, (U\Delta_j^T + \Delta_j U^T)\rangle \langle A_i U, \Delta_j\rangle + \langle A_i, UU^T - X^*\rangle \langle A_i \Delta_j, \Delta_j\rangle \right)$$

$$= 2 \sum_{i=1}^{m} \left(\langle A_i, (U\Delta_j^T + \Delta_j U^T)\rangle \langle A_i, U\Delta_j^T\rangle + \langle A_i, UU^T - X^*\rangle \langle A_i, \Delta_j \Delta_j^T\rangle \right)$$

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

$$e_j e_j^T = e_j e_j^T e_j e_j^T$$

$$\Rightarrow U \Delta_j^T = U e_j e_j^T \Delta_j^T \quad , \quad \Delta_j \Delta_j^T = (U - U^* R) e_j e_j (U - U^* R)$$

by $\nabla f(U) = 0$, i.e. $\sum_i \langle A_i, UU^T - X^* \rangle A_i U = 0$

$$\Rightarrow \sum_i \langle A_i, UU^T - X^* \rangle A_i U e_j = 0 \qquad (\langle LHS, U e_j \rangle = 0)$$

$$\Rightarrow \sum_i^m \langle A_i, UU^T - X^* \rangle \langle A_i, U e_j e_j^T U^T \rangle = 0 \quad \cdots \quad (\ast 1)$$

and $\sum_i \langle A_i, UU^T - X^* \rangle \langle A_i, U e_j e_j^T (U^* R)^T \quad \cdots \quad (\ast 2) \quad (\langle LHS, U^* R e_j \rangle = 0$

$$\Delta_j \Delta_j^T = (U - U^* R) e_j e_j^T (U - U^* R)^T$$

$$= U e_j e_j^T U^T - U e_j e_j^T (U^* R)^T - U^* R e_j e_j^T U^T + U^* R e_j e_j^T (U^* R)^T$$

and $R e_j e_j^T R^T = e_j e_j^T$, $U^* e_j e_j^T U^* = (U^* e_j e_j^T)(U^* e_j e_j^T)^T$

$$\langle A_i, UU^T - X^* \rangle \langle A_i, \Delta_j \Delta_j^T \rangle = \langle A_i, UU^T - X^* \rangle \langle A_i, U^* e_j e_j^T (U^* e_j e_j^T)^T \rangle$$

$$= - \langle A_i, UU^T - U^* U^{*T} \rangle \langle A_i, U e_j e_j^T U - U^* e_j e_j^T U^{*T}$$

$$\Rightarrow \mathrm{vec}(\Delta_j)^T [\nabla^2 f(U)] \mathrm{vec}(\Delta_j)$$

$$= \sum_{i=1}^m 4 \langle A_i, U e_j e_j \Delta_j^T \rangle^2 - 2 \langle A_i, UU^T - U^* U^{*T} \rangle \langle A_i, U e_j e_j^T U^T - U^* e_j e_j^T U^{*T} \rangle$$

$$\sum_{j=1}^r e_j e_j^T = I_r \Rightarrow$$

$$\sum_{j=1}^r \mathrm{vec}(\Delta_j)^T \nabla^2 f(U) \mathrm{vec}(\Delta_j)$$

$$= \sum_{i=1}^m \left(\sum_{j=1}^r 4 \langle A_i, U \Delta_j^T \rangle^2 - 2 \langle A_i, UU^T - U^* U^{*T} \rangle^2 \right)$$

$\square$

Corollary 4.1:

Let U be local minimum of $f(U)$

$\Rightarrow \forall$ $r\times r$ orthonormal matrix R.

$$\sum_{j=1}^{r}\sum_{i=1}^{m}\langle A_i, U\Delta_j^T\rangle^2 \geq \frac{1}{2}\sum_{i=1}^{m}\langle A_i, UU^T-U^*U^{*T}\rangle^2$$

Further, for A satisfying $(2r,\delta)$-RIP we have

$$\sum_{j=1}^{r}\|Ue_je_j^T(U-U^*R)^T\|_F^2 \geq \frac{1-\delta}{2(1+\delta)}\|UU^T-U^*U^{*T}\|_F^2$$

<pf> rank U, rank $\Delta_j \leq r$ $\Rightarrow$ rank $(U\Delta_j^T) \leq 2r$

$$\Rightarrow \sum_{i=1}^{m}\langle A_i, U\Delta_j^T\rangle^2 \leq m(1+\delta)\|U\Delta_j^T\|_F^2$$

$$\sum_{i=1}^{m}\langle A_i, UU^T-X^*\rangle^2 \geq m(1-\delta)\|UU^T-X^*\|_F^2$$

$$\Rightarrow \sum_{j=1}^{r} m(1+\delta)\|U\Delta_j^T\|_F^2 \geq \frac{1}{2}m(1-\delta)\|UU^T-U^*U^{*T}\|_F^2$$

Lemma 4.4.

$U, U^*\in\mathbb{R}^{r\times r}$, Q is orthonormal matrix that span the column space of U

$\Rightarrow \exists$ $r\times r$ orthonormal R s.t. for U 1st-order stationary pt of $f(U)$

$$\sum_{j=1}^{r}\|Ue_je_j^T(U-U^*R)^T\|_F^2 \leq \frac{1}{8}\|UU^T-U^*U^{*T}\|_F^2+\frac{34}{8}\|(UU^T-U^*U^{*T})QQ^T\|_F^2$$

<pf> omit.