CSC2414: Discrepancy Theory in Computer Science Fall 2015

Lecture 8 — November 16, 2015

Aleksandar Nikolov *Scribe: Assimakis Kattis*

γ_2 and `herdisc`: the Lower Bound

Last lecture: Recall that for $A \in \mathbb{R}^{m\times n}$, we defined the following norms:

γ_2 norm:

$$\gamma_2(A) = \min\{r(U)\cdot c(V) : UV = A\}$$

Norms of Rows: For a_{i*} the i-th row of A:

$$r(A) = \max_{i=1}^{m} \|a_{i*}\|_2$$

Norms of Columns: For a_{*i} the i-th column of A:

$$c(A) = \max_{i=1}^{n} \|a_{*i}\|_2$$

Theorem 1 (Larsen)**.** *For all $A \in \mathbb{R}^{m\times n}$:*

$$\text{herdisc}\,(A) = \gamma_2(A)\cdot O(\sqrt{\log m})$$

This lecture: We show the following result:

Theorem 2. *For all $A \in \mathbb{R}^{m\times n}$:*

$$\text{herdisc}\,(A) = \gamma_2(A)\cdot \Omega\left(\frac{1}{\log \text{rank}\, A}\right)$$

We have previously shown that the determinant lower bound, given by:

$$\text{detlb}\,(A) = \max_{k=1}^{\min(m,n)} \max_{\substack{S\subseteq[m]\\ T\subseteq[n]\\ |S|=|T|=k}} |\det A_{S,T}|^{1/k}$$

where $A_{S,T}$ the subset of A indexed by S and T, satisfies:

$$\text{herdisc}\,(A) \geq \frac{1}{2}\,\text{detlb}\,(A).$$

Thus, it suffices to show that:

Theorem 3 ([MNT14]). *For all $A \in \mathbb{R}^{m\times n}$:*

$$\gamma_2(A) \geq \operatorname{detlb}(A)$$

$$\operatorname{detlb}(A) = \gamma_2(A) \cdot \Omega\left(\frac{1}{\log \operatorname{rank} A}\right)$$

Dual Characterization of γ_2: We have shown that the following vector program has $\gamma_2(A)$ as its optimum:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \langle u_i, v_j\rangle = A_{ij} \\ & \langle u_i, u_i\rangle \leq t \\ & \langle v_j, v_j\rangle \leq t \\ & u_i, v_j \in \mathbb{R}^{m+n} \\ \text{where} & (i,j) \in [m]\times[n], \end{array}$$

This exhibits strong duality, meaning that it is equal to its dual (maximization) problem, which is shown below:

$$\begin{array}{ll} \text{maximize} & \|B\|_{tr} \\ \text{subject to} & B_{ij} = p_i q_j A_{ij} \\ & \displaystyle\sum_{i=1}^{m} p_i^2 = \sum_{j=1}^{n} q_j^2 = 1 \\ & p_i, q_j \geq 0 \\ \text{where} & (i,j) \in [m]\times[n], \end{array}$$

where $\|B\|_{tr}$ is the trace or nuclear norm, which is equal to the sum of its singular values:

$$\|B\|_{tr} = \sum_{i=1}^{\min(m,n)} \sigma_i.$$

Claim. $\gamma_2(A) \geq \operatorname{detlb}(A)$

Proof. Pick optimal S, T with $S \subseteq [m], T \subseteq [n]$ and $|S| = |T| = k$ for which $|\det A_{S,T}|^{1/k} = \operatorname{detlb}(A)$. Suffices to show that:

$$\gamma_2(A) \geq |\det A_{S,T}|^{1/k}$$

For the dual maximization problem for γ_2, define the following:

$$p_i = \begin{cases} 1/\sqrt{k} & i \in S \\ 0 & \text{otherwise} \end{cases}$$

$$q_j = \begin{cases} 1/\sqrt{k} & j \in T \\ 0 & \text{otherwise} \end{cases}$$

Since the above are orthogonal, we can reorder the corresponding matrix $B = p_i q_j A_{ij}$:

$$B' = \begin{bmatrix} \frac{1}{k} A_{S,T} & 0 \\ 0 & 0 \end{bmatrix}$$

from which we get that:

$$\|B\|_{tr} = \|B'\|_{tr} = \frac{1}{k}\|A_{S,T}\|_{tr}$$

By the AM-GM inequality and since $\gamma_2(A)$ is the maximum possible value for $\|B\|_{tr}$, this implies:

$$\gamma_2(A) \geq \frac{1}{k}\|A_{S,T}\|_{tr} \geq |\det A_{S,T}|^{1/k}$$

□

Bucketing Lemma: For all $\sigma \in \mathbb{R}_+^r$, $\exists R \in [r]$ for which:

$$\sum_{i \in R} \sigma_i \geq \frac{1}{2 \log 2r} \sum_{i=1}^{r} \sigma_i$$

$$\forall i, j \in R, \sigma_i \leq 2\sigma_j$$

Proof. Without loss of generality, assume that $\sum_{i=1}^r \sigma_i = 1$. Then define the following sets, where $1 \leq k \leq \lceil \log 2r \rceil$:

$$R_k = \{i : (1/2)^{k-1} \geq \sigma_i \geq (1/2)^k\}$$
$$R_\infty = \{i : \sigma_i \leq (1/2r)\}$$

The main motivation behind this construction is that we can ignore R_∞, and the rest will follow by averaging. Note also that all R_k satisfy the second property.

Since $|R_\infty| < r$, we have that:

$$\sum_{i \in R_\infty} \sigma_i < |R_\infty| \cdot \frac{1}{2r} < \frac{1}{2}$$

$$\sum_{k=1}^{\log 2r} \sum_{i \in R_k} \sigma_i = 1 - \sum_{i \in R_\infty} \sigma_i > 1/2$$

This means that as we have $\log 2r$ terms, we get:

$$\log 2r \cdot \min_k \sum_{i\in R_k} \sigma_i > 1/2$$

Thus, $\exists R_l$ for which:

$$\sum_{i\in R_l} \sigma_i \geq \frac{1}{2\log 2r}$$

□

We can immediately see that for such a set R, we have:

Corollary 4. *For R satisfying the conditions of the Bucketing Lemma:*

$$\frac{1}{|R|}\sum_{i\in R} \sigma_i \leq 2\left(\prod_{i\in R}\sigma_i\right)^{1/|R|}$$

Proof.

$$\frac{1}{|R|}\sum_{i\in R} \sigma_i = \max_{i\in R}\sigma_i \leq 2\min_{i\in R}\sigma_i \leq 2\left(\prod_{i\in R}\sigma_i\right)^{1/|R|}$$

□

Claim 5.

$$\operatorname{detlb}(A) = \gamma_2(A)\cdot\Omega\left(\frac{1}{\log \operatorname{rank} A}\right)$$

Proof. Take a feasible solution (B, p, q) to the dual maximization problem for $\gamma_2(A)$. This implies that $\gamma_2(A) = \|B\|_{tr}$.

Now, let the singular value decomposition (SVD) of B be $B = U\Sigma V^T$. Here, $r = \operatorname{rank} B$, $U \in \mathbb{R}^{m\times r}$, $V \in \mathbb{R}^{n\times r}$, $U^TU = I$ and Σ a diagonal matrix with the singular values of B on the diagonal.

Pick $R \in [n]$ with $|R| = k$ to be a subset of singular values that satisfies the conditions of the Bucketing lemma. If we define $C := U_R^T B$, where U_R the subset of U indexed by R, then the singular values of C are $\{\sigma_i\}_{i\in R}$. This means that:

$$\left|\det CC^T\right|^{\frac{1}{2k}} = \left|\prod_{i\in R}\sigma_i\right|^{\frac{1}{k}} \geq \frac{1}{2k}\sum_{i\in R}\sigma_i \geq \frac{1}{4k\log 2r}\sum_{i=1}^{r}\sigma_i = \frac{1}{4k\log 2r}\|B\|_{tr} \tag{1}$$

Cauchy-Binet Formula: For $X, Y \in \mathbb{R}^{m\times n}$:

$$\det XY^T = \sum_{\substack{S\subseteq[n]\\|S|=m}} \det X_S \det Y_S$$

If we define $P \in \mathbb{R}^{m\times m}$ and $Q \in \mathbb{R}^{n\times n}$ as the diagonal matrices with p_i and q_j as the diagonal entries respectively, then we have $B = PAQ$. Similarly, $C = U_R^T B = U_R^T PAQ$.

Define $D := U_R^T PA$ so that $C = DQ$. By applying Cauchy-Binet to $C \in \mathbb{R}^{k\times n}$ we get:

$$\det(CC^T) = \sum_{\substack{S\subseteq[n]\\|S|=k}} \det C_S \det C_S = \sum_{\substack{S\subseteq[n]\\|S|=k}} (\det C_S)^2 = \sum_{\substack{S\subseteq[n]\\|S|=k}} (\det D_S Q_S)^2$$

$$= \sum_{\substack{S\subseteq[n]\\|S|=k}} (\det D_S)^2 \left(\prod_{j\in S} q_j^2\right) \leq \left(\max_{\substack{S\subseteq[n]\\|S|=k}} (\det D_S)^2\right)\left(\sum_{\substack{S\subseteq[n]\\|S|=k}} \prod_{j\in S} q_j^2\right)$$

which follows by Hölder.

By picking distinct j from each of the k sums, we will get each j $k!$ times. Therefore, this implies:

$$\sum_{\substack{S\subseteq[n]\\|S|=k}} \prod_{j\in S} q_j^2 \leq \frac{1}{k!}\left(\sum_{j=1}^{n} q_j^2\right)^k = \frac{1}{k!}$$

Thus, $\exists S \subseteq [n]$ such that:

$$(\det D_S)^{1/k} \geq (k!)^{1/2k} \cdot (\det CC^T)^{1/2k}$$

or equivalently:

$$\max_{\substack{S\subseteq[n]\\|S|=k}} |\det D_S|^{1/k} \geq (k!)^{1/2k} \cdot (\det CC^T)^{1/2k}$$

which by Stirling means:

$$(k!)^{1/2k} \cdot (\det CC^T)^{1/2k} \geq \sqrt{\frac{k}{e}}(\det CC^T)^{1/2k} = \Omega(\sqrt{k})(\det CC^T)^{1/2k}$$

Thus, by applying equation (1) here this implies:

$$(\det D_S)^{1/k} \geq \frac{\|B\|_{tr}}{4e\sqrt{k}\log 2r} \tag{2}$$

Consider the orthonormal matrix $W \in \mathbb{R}^{m\times m}$ for which the first k columns are equal to the columns of U_R. Such a matrix always exists since we can complete the orthonormal basis for $\mathbb{R}^m$ starting with the column vectors of U_R. The $m - k$ new vectors we get can be used to define the rest of the columns of W.

Define $E_S := PA_S \in \mathbb{R}^{m\times k}$, meaning that $D_S = U_R^T E_S$. It can be shown that:

$$\det(E_S^T E_S) = \det((E_S^T W)(W^T E_S)) = \det((E_S^T W)(E_S^T W)^T)$$

$$= \sum_{\substack{T\subseteq[n]\\|T|=k}} \det((E_S^T W)_T)^2 = \sum_{\substack{T\subseteq[n]\\|T|=k}} \det(E_S^T W_T)^2 \geq \det(E_S^T U_R)^2 = \det(U_R^T E_S)^2$$

$$\therefore \det(E_S^T E_S) \geq \det(D_S)^2$$

Now, we can apply the exact same analysis as in (2), but this time to $D_S^T = (A_S)^T P$ instead of C. This means that $\exists T \in [m]$ for which:

$$\max_{\substack{T\subseteq[m]\\|T|=k}} (\det A_{S,T})^{1/k} \geq (k!)^{1/2k} \cdot \det\left(A_S^T P^2 A_S\right)^{1/2k} = (k!)^{1/2k} \cdot \det\left(E_S^T E_S\right)^{1/2k}$$

Putting all of this together and applying Stirling just like before, we get that:

$$\max_{\substack{S\subseteq[n]\\T\subseteq[m]\\|S|=|T|=k}} |\det A_{S,T}|^{1/k} \geq \frac{\|B\|_{tr}}{4e\log(2r)}$$

By maximizing over all k, this yields the desired result.

□

References

[MNT14] J. Matousek, A. Nikolov, and K. Talwar. Factorization Norms and Hereditary Discrepancy. *ArXiv e-prints*, August 2014.