CSC2414: Discrepancy Theory in Computer Science	Fall 2015

Lecture 9 — November 23, 2015

Aleksandar Nikolov — *Scribe: Nick Spooner*

1 Properties of γ_2

Recall that $\gamma_2(A)$ is defined for $A \in \mathbb{R}^{m\times n}$ as follows:

$$\gamma_2(A) = \min\{r(U)\cdot c(V) : UV = A, U \in \mathbb{R}^{m\times k}, V \in \mathbb{R}^{k\times n}, k \in \mathbb{N}\} \ ,$$

where $r(U)$ is the maximum row norm of U, and $c(V)$ is the maximum column norm of V. We showed in the previous lecture that there exists a constant C such that

$$\frac{\gamma_2(A)}{C\log \operatorname{rk} A} \le \operatorname{herdisc}(A) \le C\sqrt{\log m}\cdot\gamma_2(A) \ . \tag{1}$$

We give some other useful properties of γ_2.

1. *Monotonicity.* $\gamma_2(A_{S,T}) \le \gamma_2(A)$, where $A_{S,T}$ is the submatrix of A whose rows are indexed by $S \subseteq [m]$ and whose columns are indexed by $T \subseteq [n]$.
2. *Transpose.* $\gamma_2(A) = \gamma_2(A^T)$, where A^T is the matrix transpose of A.
3. *Diagonal block matrices.* $\gamma_2\left(\begin{pmatrix} A & 0 \\ 0 & B\end{pmatrix}\right) = \max(\gamma_2(A), \gamma_2(B))$.
4. *Triangle inequality.* $\gamma_2(A+B) \le \gamma_2(A) + \gamma_2(B)$.
5. *Union.* $\gamma_2\left(\begin{pmatrix} A & B\end{pmatrix}\right) \le \sqrt{\gamma_2(A)^2 + \gamma_2(B)^2}$.

Most of these properties follow straightforwardly from the definitions. We give a detailed proof of Property 4.

Proof of triangle inequality. Let U_A, V_A be such that $U_AV_A = A$, and $r(U_A) = c(V_A) = \sqrt{\gamma_2(A)}$. This can always be achieved simply by scaling the matrices appropriately. We take U_B, V_B similarly. Let $U := \begin{pmatrix} U_A & U_B\end{pmatrix}$, and $V := \begin{pmatrix} V_A \\ V_B\end{pmatrix}$. Then clearly $UV = A + B$. Moreover

$$\begin{aligned} r(U)^2 &= \max_{i=1}^m \|U_{i*}\|_2^2 = \max_{i=1}^m \left(\|(U_A)_{i*}\|_2^2 + \|(U_B)_{i*}\|_2^2\right) \\ &\le \max_{i=1}^m \|(U_A)_{i*}\|_2^2 + \max_{i=1}^m \|(U_B)_{i*}\|_2^2 = r(U_A)^2 + r(U_B)^2 = \gamma_2(A) + \gamma_2(B) \ . \end{aligned}$$

The same inequality holds for $c(V)^2$, so

$$\gamma_2(A+B) \le \sqrt{r(A)^2\cdot c(V)^2} \le \gamma_2(A) + \gamma_2(B) \ . \qquad \square$$

Remark 1. Using the bounds in (1), we can obtain approximate versions of the above properties for herdisc.

1.1 Kronecker products

For matrices $A \in \mathbb{R}^{p\times q}$, $B \in \mathbb{R}^{r\times s}$, the Kronecker (tensor) product $A \otimes B \in \mathbb{R}^{pr\times qs}$ is given by the block matrix

$$A \otimes B = \begin{pmatrix} A_{11}\cdot B & A_{12}\cdot B & \cdots \\ A_{21}\cdot B & A_{22}\cdot B & \\ \vdots & & \ddots \end{pmatrix} .$$

Lemma 2 (Property 6)**.** *$\gamma_2(A \otimes B) = \gamma_2(A)\gamma_2(B)$.*

Remark 3. This property does not hold for the combinatorial discrepancy disc(A).

Proof. We make use of a basic property of the tensor product: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. Applying this property to the singular value decompositions of A and B, we see that if A has singular values $\sigma_1, \ldots, \sigma_p$ and B has singular values $\tau_1, \ldots, \tau_r$ then $A \otimes B$ has singular values $\sigma_1\tau_1, \ldots, \sigma_1\tau_r, \ldots, \sigma_p\tau_1, \ldots, \sigma_p\tau_r$.

First we prove that $\gamma_2(A \otimes B) \le \gamma_2(A)\gamma_2(B)$. We take U_A, V_A such that $A = U_AV_A$ and $r(U_A)c(V_A) = \gamma_2(A)$; similarly we have U_B, V_B. Let $U := U_A \otimes U_B$ and $V := V_A \otimes V_B$. Then $UV = A\otimes B$, by the basic property of tensor products. Moreover we have that $r(U) = r(U_A)r(U_B)$, since the rows of U have the form $u_A \otimes u_B$ where u_A is a row of A and u_B is a row of B, and for any two vectors u_1, u_2, $\|u_1 \otimes u_2\|_2 = \|u_1\|_2 \cdot \|u_2\|_2$. The same property holds of the columns of V and hence of $c(V)$, and so

$$\gamma_2(A \otimes B) \le r(U)c(V) = (r(U_A)c(V_A)) \cdot (r(U_B)c(V_B)) = \gamma_2(A)\gamma_2(B) \ .$$

It remains to show that $\gamma_2(A\otimes B) \ge \gamma_2(A)\gamma_2(B)$. For this we make use of the dual of the semidefinite program for computing γ_2. By strong duality, it holds that (cf. last lecture)

$$\gamma_2(A) = \max\{\|PAQ\|_{\mathrm{tr}} : P, Q \text{ nonnegative diagonal matrices s.t. } \mathrm{tr}(P^2) = \mathrm{tr}(Q^2) = 1\} \ .$$

Let P_A, Q_A be such that $\|P_AAQ_A\|_{\mathrm{tr}} = \gamma_2(A)$, and P_B, Q_B likewise. Let $P := P_A \otimes P_B$ and $Q = Q_A \otimes Q_B$. Note that $\mathrm{tr}((P_A \otimes P_B)^2) = \mathrm{tr}(P_A^2 \otimes P_B^2) = \mathrm{tr}(P_A^2)\,\mathrm{tr}(P_B^2) = 1$ by easy properties of the Kronecker product, and the same holds for Q, hence P, Q is a feasible solution. Finally from the properties of the singular values of Kronecker products, we get

$$\begin{aligned}\|P(A \otimes B)Q\|_{\mathrm{tr}} &= \|(P_AAQ_A) \otimes (P_BBQ_B)\|_{\mathrm{tr}} = \sum_i\sum_j \sigma_i\tau_j = (\sum_i \sigma_i)(\sum_j \tau_j) \\ &= \|P_AAQ_A\|_{\mathrm{tr}} \cdot \|P_BBQ_B\|_{\mathrm{tr}} = \gamma_2(A) \cdot \gamma_2(B) \ .\end{aligned}$$

□

2 Discrepancy of corners

Recall from Lecture 1 that for $y \in \mathbb{R}^d$ we define the corner

$$\mathcal{C}(y) := \{x \in \mathbb{R}^d : 0 \le x_i \le y_i, i = 1, \ldots, d\} \ .$$

For $d \in \mathbb{N}$ the set $\mathcal{C}_d := \{\mathcal{C}(y) : y \in [0, 1]^d\}$. Let P be a finite subset of $[0, 1]^d$; then $\mathcal{C}_d\,|_P$ is the set of subsets of P of the form $C(y) \cap P$ for some $y \in [0, 1]^d$. We define the combinatorial discrepancy of $\mathcal{C}_d$, $\mathrm{disc}(n, \mathcal{C}_d) := \sup_P \mathrm{disc}(\mathcal{C}_d|_P)$, where the supremum is taken over subsets $P \subseteq [0,1]^d$ of size n. Recall also that the continuous discrepancy $D(n, \mathcal{C}_d) \leq O(1) \cdot \mathrm{disc}(n, \mathcal{C}_d)$.

Open problem. We know that (approximately) $\log^{(d-1)/2} n \leq D(n, \mathcal{C}_d) \leq \log^{d-1} n$. Can we get a tighter bound?

The following theorem suggests that better bounds for $D(n, \mathcal{C}_d)$ are unlikely to come from better bounds for $\mathrm{disc}(n, \mathcal{C}_d)$.

Theorem 4. *For $d \in \mathbb{N}$, it holds that*

$$\Omega(\log^{d-1} n) \leq \mathrm{disc}(n, \mathcal{C}_d) \leq O(\log^{d+1/2} n) \ .$$

Proof sketch. Let $Q := [n]^d \subseteq [0,n]^d$ be the set of d-vectors whose coordinates are positive integers at most n. (Note: we can scale this set into $[0,1]^d$ so that it fits the definitions above.) Let $\mathcal{S} := \mathcal{C}_d|_Q$. Then $\mathcal{S} := \{[y_1] \times \ldots \times [y_d] : y_1, \ldots, y_d \in [n]\}$, and the incidence matrix A of $\mathcal{S}$ is $T_n^{\otimes d}$, the d-wise Kronecker product of the $n \times n$ lower triangular matrix with 1s below (and on) the diagonal.

Proposition 5. $\gamma_2(T_n) = \Theta(\log n)$.

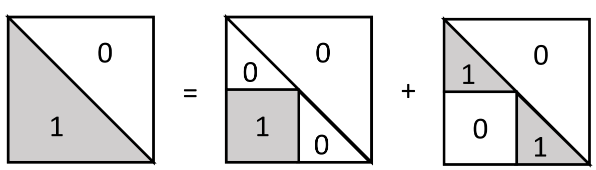
Given the proposition, it is not too difficult to show that the upper and lower bounds hold.

Lower bound. By Lemma 2, $\gamma_2(A) = \Theta(\log^d n)$. Then inequality (1) gives that $\mathrm{herdisc}(A) \geq \Omega(\gamma_2(A)/\log \mathrm{rk}\, A) = \Omega(\log^{d-1} n)$. By the definition of hereditary discrepancy, there exists a subset $P \subseteq Q$, $|P| = n$, such that the discrepancy $\mathrm{disc}(\mathcal{S}|_P) = \Omega(\log^{d-1} n)$. Since $\mathcal{S} = \mathcal{C}_d|_Q$, and $P \subseteq Q$, clearly $\mathcal{S}|_P = \mathcal{C}_d|_P$, so $\mathrm{disc}(n, \mathcal{C}_d) \geq \mathrm{disc}(\mathcal{S}|_P) = \Omega(\log^{d-1} n)$.

Upper bound. We may assume $P \subseteq [n]^d$, since for any $P \subseteq [0,n]^d$ we can transform it into $P' \subseteq [n]^d$ such that $|P'| = |P|$ and $\mathrm{disc}(\mathcal{C}_d|_P) \leq \mathrm{disc}(\mathcal{C}_d|_{P'})$. Then $\mathrm{disc}(n, \mathcal{C}_d) \leq \mathrm{herdisc}(\mathcal{S})$. The γ_2 upper bound gives $\mathrm{herdisc}(\mathcal{S}) = O(\sqrt{\log n^d})\gamma_2(A) = O(\log^{d+1/2} n)$, which concludes the proof. $\square$

Proof of Proposition 5. We show the upper and lower bounds separately.

1. $\gamma_2(T_n) = O(\log n)$. Notice that we can decompose T_n as follows:



Or, written as a sum of block matrices,

$$T_n = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} T_{n/2} & 0 \\ 0 & T_{n/2} \end{pmatrix}.$$

Note that $\gamma_2\left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right) = 1$. From Property 3 of γ_2, $\gamma_2\left(\begin{pmatrix} T_{n/2} & 0 \\ 0 & T_{n/2} \end{pmatrix}\right) = \gamma_2(T_{n/2})$. Thus $\gamma_2(T_n) \leq 1 + \gamma_2(T_{n/2})$ by the triangle inequality, and solving the recurrence gives the upper bound.

2. $\gamma_2(T_n) = \Omega(\log n)$. Here we once again make use of the dual. By the normality conditions on P and Q, for any matrix $A \in \mathbb{R}^{k\times k}$ we have that $P = Q = \frac{1}{\sqrt{k}} I_k$ is a feasible solution, and so $\gamma_2(A) \geq \|PAQ\|_{\text{tr}} = \frac{1}{k}\|A\|_{\text{tr}}$. Let $B := \begin{pmatrix} T_n & T_n^T \\ T_n^T & T_n \end{pmatrix}$. It is not difficult to see that this is a circulant matrix, i.e. each column is the previous column rotated by one row. The eigenvalues of such a matrix are the DFT coefficients of its first column, and so the i-th singular value of B is approximately n/i, so $\|B\|_{\text{tr}} = \Theta(n \log n)$. Then finally

$$\gamma_2(T_n) \geq \frac{1}{4}\gamma_2(B) \geq \frac{1}{8n}\|B\|_{\text{tr}} = \Omega(\log n) \ ,$$

which proves the claim. □

3 Data structure lower bounds

3.1 Range counting

Let $d \in \mathbb{N}$, point set $P \subseteq \mathbb{R}^d$, and weight function $w : P \to \mathbb{Z}$. We are interested in designing a data structure which supports two operations:

Update. Given a pair $(p, x) \in P \times \mathbb{Z}$, set $w(p) := w(p) + x$.

Query. Given $z \in \mathbb{R}^d$, return $\sum_{p \in \mathcal{C}(z)|_P} w(p)$.

3.2 The oblivious group model

We define a restricted model of a data structure. In this model, the data structure retains s values $y := (y_1, \ldots, y_s)$ where each $y_i \in \mathbb{R}$ is a linear combination of the $w(p)$ for $p \in P$. Let U, V be such that $UV = A$, where A is the incidence matrix of $\mathcal{C}_d\,|_P$. V encodes the linear combinations of group elements which are used to compute y: $y = Vw$. U encodes the linear combinations of y_i which are used to answer queries; the condition $UV = A$ is necessary for correctness. Then our operations are constrained to be of the following form:

Update. Given a pair (p, x), $y := y + xV_{*p}$, where V_{*p} is the p-th column of V.

Query. Given z, return $\langle U_{i*}, y\rangle$, where the i-th row of A is the indicator vector of the set $\mathcal{C}(z) \cap P$.

The time complexity of updates, $t_u := \max_p \operatorname{nnz}(V_{*p})$, where for a vector u, $\operatorname{nnz}(u)$ is defined as the number of non-zero entries in u. Similarly the time complexity of queries $t_q := \max_z \operatorname{nnz}(U_{z*})$.

The one-dimensional case is well-understood.

Theorem 6 (Fredman '82 [1])**.** *For $d = 1$, $t_u + t_q = \Omega(\log n)$.*

For higher dimensions, we require an additional parameter Δ, a bound on the absolute values of entries in U, V.

Theorem 7 (Larsen '11 [2])**.** *For $d \in \mathbb{N}$, $t_u t_q \geq d^2\Omega(\log^2 n)/\Delta^4$.*

Proof sketch. It is not hard to see that $r(U) \leq \Delta\sqrt{t_q}$, and $c(V) \leq \Delta\sqrt{t_u}$. Then $\gamma_2(A) \leq \Delta^2\sqrt{t_u t_q}$, so $\sqrt{t_u t_q} \geq \frac{\Omega(\log n^d)}{\Delta^2}$. $\square$

For many natural data structures, $\Delta = 1$. Indeed for any $d \in \mathbb{N}$ there exists a data structure with $\Delta = 1$ which matches the lower bound. In the case $d = 1$, the Fredman bound shows that the dependence on Δ is not required, but for $d > 1$ it remains open whether larger values of Δ allow for more efficient data structures.

References

[1] Michael L. Fredman. 1982. The Complexity of Maintaining an Array and Computing Its Partial Sums. J. ACM 29, 1 (January 1982), 250-260.

[2] Larsen, K.G. 2011. On Range Searching in the Group Model and Combinatorial Discrepancy. IEEE 52nd Annual Symposium on the Foundations of Computer Science (FOCS), pp.542-549.