Bichromatic Combinatorial Discrepancy

Jeff M. Phillips

September 12, 2005

Consider a set of n red points R, and a set of n blue points B which lie in the plane. Let $B(\mathcal{R})$ be the number of blue points in a region $\mathcal{R}$, and similarly let $R(\mathcal{R})$ be the number of red points in $\mathcal{R}$. The discrepancy of any region $\mathcal{R}$ is defined as

$$\Delta_{\mathcal{R}} = \mid R(\mathcal{R}) - B(\mathcal{R}) \mid,$$

the difference between the number of red points and the number of blue points in that region. In 1995, Dobkin, Gunopulos, and Maass [1], showed how to compute the axis-aligned rectangle with the maximum bichromatic combinatorial discrepancy in time $O(n^2 \log n)$.

This is an important low-level operation in machine learning and computer graphics. This technique surprisingly does not extend to other shapes such as axis-aligned squares, but does extend to other discrepancy measures such as numerical discrepancy. Consider a set of points S where for all $s \in S$, $s \in [0,1] \times [0,1]$. Numerical discrepancy is defined over a region $\mathcal{R}$ with area $|\mathcal{R}|$ as

$$\mathcal{D}_{\mathcal{R}} = \mid |\mathcal{R}| - S(\mathcal{R})/n \mid .$$

I will present this, clever, yet simple algorithm.

References

[1] David P. Dobkin, Dimitrious Gunopulos, and Wolfgang Maass. Computing the Maximum Bichromatic Discrepancy, with applications to Computer Graphics and Machine Learning. *Neural and Computational Learning*, 1995.