Majorization and network problems

Geir Dahl*

*Centre of Mathematics for Applications and Department of Informatics, University of Oslo
P.O. Box 1053 Blindern, NO-0316 Oslo, NORWAY. (geird@math.uio.no)

Abstract

The notion of *majorization* has its roots in matrix theory and economics. Loosely speaking, for two vectors $x, y \in \mathbb{R}$ satisfying $\sum_i x_i = \sum_i y_i$, we say that x is majorized by y if the components in x are "less spread out" than the components in y. This may be expressed in terms of linear inequalities for the partial sums in these vectors (see below). This notion arises in a wide range of contexts in mathematical areas, e.g. in combinatorics, probability, matrix theory, numerical analysis and quantum physics.

The purpose of this presentation is to give a brief introduction to majorization theory and present some recent developments where majorization is studied in connection with familiar network structures (trees and transportation matrices).

Keywords: *Majorization, trees, distances, transportation matrices.*

1 The majorization ordering

We present the basic notion and mention some classical results. For $x \in \mathbb{R}^n$ we let $x_{[j]}$ denote the jth largest number among the components of a vector x. If $x, y \in \mathbb{R}^n$ we say that x is *weakly majorized* by y, denoted by $x \preceq_* y$, provided that

$$\sum_{j=1}^{k} x_{[j]} \leq \sum_{j=1}^{k} y_{[j]} \quad (k = 1, 2, \ldots, n)$$

If, in addition, $\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j$ holds, then x is *majorized* by y, denoted by $x \preceq y$. We refer to Marshall and Olkin's book [9] for a comprehensive study of majorization and its role in many branches of mathematics and applications. Another useful reference here on this is [1]. As an example we have

$$(\frac{1}{n}, \ldots, \frac{1}{n}) \preceq (\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \preceq (1, 0, \ldots, 0).$$

We also define the majorization-like ordering $\preceq'_*$: for vectors $x, y \in \mathbb{R}^n$ we write $x \preceq'_* y$ whenever $\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j$ for $k = 1, \ldots, n$. Majorization turns out to be an underlying structure for several classes of inequalities. One such simple example is the arithmetic-geometric mean inequality. Another example is a majorization order between the diagonal entries and the eigenvalues of a real symmetric matrix. Actually, several interesting inequalities arise by applying some ordering-preserving function to a suitable majorization ordering.

The following theorem contains some important classical results concerning majorization, due to Hardy, Littlewood, Polya (1929) and Schur (1923). Recall that a (square) matrix is *doubly stochastic* if it is nonnegative and all row and column sums are equal to one.

Theorem 1 *Let $x, y \in \mathbb{R}^n$. Then the following statements are equivalent.*

(i) $x \preceq y$.

(ii) *There is a doubly stochastic matrix A such that $x = Ay$.*

(iii) *The inequality $\sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i)$ holds for all convex functions $g : \mathbb{R} \rightarrow \mathbb{R}$.*

So, in particular, this shows that there is a close connection between majorization and doubly stochastic matrices. And this matrix class is closely tied to matching theory; each doubly stochastic matrix corresponds to a fractional perfect matching in a bipartite graph.

Majorization arises in an important area in combinatorics. The *Gale-Ryser theorem* characterizes the existence of a $(0,1)$-matrix with given row and column sum vectors. This condition is precisely a majorization ordering $S \preceq R^*$ where S is the column sum vector and R^* is the conjugate of the row sum vector R. This theorem may be restated in terms of bipartite graphs with given degrees. For further on this and related results, we refer to Brualdi's recent book [2].

We next briefly discuss some recent developments where majorization plays a role in some network related problems.

2 Majorization in trees

In facility location analysis the distribution of the distances between a facility and the customers may be important. Erkut [7] studied several different measures of inequality, or equity, of the distribution $d = (d_i(x) : i \leq n)$ of distances between a facility x and its customers $i = 1, 2, \ldots, n$. An example is the measure $(\sum_i (d_i - \bar{d})^2)^{1/2}$ (proportional to the standard deviation). One may then look for a facility location which minimizes the selected inequality measure since this means, in some sense, that the distances to the different customers are "as equal as possible".

Motivated by this we consider majorization in trees as introduced and discussed in [5]. Let $T = (V, E)$ be a tree with n vertices and let $d(u, v)$ denote the distance between two vertices u and v in T (so $d(u, v)$ is the number of edges in the unique uv-path in T). For $u \in V$ we define the *distance vector of* u as follows: $d(u, \cdot) = (d(u, v) : v \in V) \in \mathbf{Z}_+^n$.

We now define our main concepts. Let v_1, v_2 be two vertices in the tree T. If $d(v_1, \cdot) \preceq_* d(v_2, \cdot)$ we say that v_1 is *majorized* by v_2 and write $v_1 \preceq_* v_2$. Moreover, we say that v_1 and v_2 are *majorization-equivalent* if $v_1 \preceq_* v_2$ and $v_2 \preceq_* v_1$; this means that $d(v_1, \cdot)$ is a permutation of $d(v_2, \cdot)$. There is at most one pair of adjacent majorization-equivalent vertices. If a tree T contains two adjacent majorization-equivalent vertices, we say that T is *m-symmetric*.

Some main results on this notion of majorization are stated next. Consider a tree T and two adjacent vertices v_1 and v_2 in T. By removing the edge $[v_1, v_2]$ one gets two subtrees $T(v_1; v_2)$ and $T(v_2; v_1)$ where $T(v_1; v_2)$ contains v_1 and $T(v_2; v_1)$ contains v_2. The vertex sets of these two subtrees are denoted by $V(v_1; v_2)$ and $V(v_2; v_1)$, respectively. Let $V_1 = V(v_1; v_2)$ and $V_2 = V(v_2; v_1)$. Define the n-vector

$$N(v_1; v_2) = (N_{n-1}(v_1; v_2), N_{n-2}(v_1; v_2), \ldots, N_0(v_1; v_2))$$

where $N_k(v_1; v_2)$ equals the number of vertices in $T(v_1; v_2)$ with distance k to v_1 $(0 \leq k \leq n-1)$.

Theorem 2 *Consider a tree T and two adjacent vertices v_1 and v_2 in T (with the notation as above). Then the following statements are equivalent:*

(i) $v_1 \preceq_* v_2$

(ii) $|V_1| \geq |V_2|$ *and* $d_{V_1}(v_1, \cdot)_{[j]} \geq d_{V_2}(v_2, \cdot)_{[j]}$ *for all* $1 \leq j \leq |V_2|$

(iii) $N(v_2; v_1) \preceq_*' N(v_1; v_2)$.

Example. In the tree T in Fig. 1 the arrows on the edges indicate majorizations between the distance vectors of adjacent vertices. For instance, $u \prec_* w \prec_* x$ and $v \prec_* u$ and $u \prec_* v$. Note that majorizations hold along each path from a leaf toward the nearest vertex among u and v.

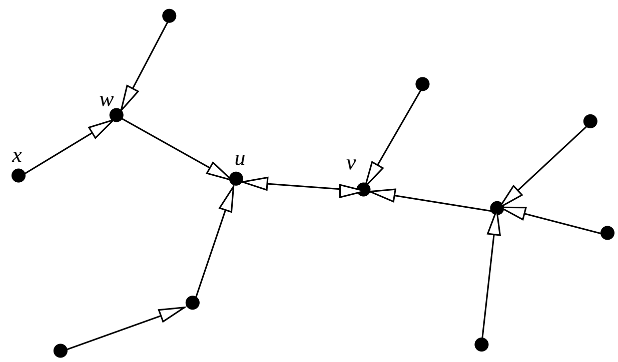

Figure 1: A tree and adjacent majorizations.

Corollary 3 *Let T be a tree and let v_1, v_2 be adjacent vertices with $v_1 \preceq_* v_2$. Then*

$$v_1 \preceq_* v_2 \preceq_* v \quad \textit{for all } v \in V(v_2; v_1)$$

i.e., for each path $v_1, v_2, v_3, \ldots, v_N$ in T where v_N is a leaf, the majorizations $v_1 \preceq_* v_2 \preceq_* v_3 \preceq_* \cdots \preceq_* v_N$ hold.

If T is m-symmetric, say that v_1 and v_2 are the adjacent majorization-equivalent vertices, then each vertex $v \in T$ majorizes both v_1 and v_2.

Furthermore, one may introduce a new center concept in trees, called the *majorization-center*, in the following way. It is a vertex set, denoted by M_T, and

(i) If T is m-symmetric, then $M_T = \{v_1, v_2\}$ where v_1 and v_2 are the two adjacent majorization-equivalent vertices.

(ii) If T is not m-symmetric, then we let M_T be the intersection of all vertex sets $V(v_1; v_2)$ taken over all adjacent vertices v_1 and v_2 for which the majorization $v_1 \preceq_* v_2$ holds.

One may observe that the majorization-center M_T is a subtree of T; this is a consequence of Corollary 3. In fact, if v_1 and v_2 are adjacent and $v_1 \preceq_* v_2$ holds, then we can eliminate all vertices in $V(v_2; v_1)$ from the majorization-center (except v_2 when T is m-symmetric). This gives an elimination procedure for finding M_T.

The next result collects some further properties of the majorization-center.

Theorem 4 *The majorization-center M_T is a subtree of T with the following properties:*

(i) *M_T contains no leaf of T except when $|V| \leq 2$.*

(ii) *M_T contains every center vertex and every median of T.*

(iii) *If T is an m-symmetric tree, then $M_T = C_T$, where C_T is the set of center vertices of T.*

We conclude this section with some additional remarks:

- The majorization-center of a tree T may be calculated efficiently by a simple algorithm based on Theorem 2. For each edge $[v_1, v_2]$ in T one calculates the two vectors $N(v_1; v_2)$ and $N(v_2; v_1)$; this is done in $O(n)$ steps using breadth-first-search in T (in the two associated subtrees). Then one checks the majorization condition (iii) in Theorem 2 (requires $O(n)$ additions/subtractions). This is an $O(n^2)$ algorithm for finding the majorization-center M_T.

- Interestingly, one may show that any given tree T_0 may arise as the majorization-center of a suitably chosen tree T. However, computational experiments done by generating large classes of "random" trees, suggest that the majorization-center typically is very small, usually it contains at most three vertices.

- The results above may be used to introduce interesting classes of functions defined on the distance vectors of vertices in a tree T such that a minimum of such functions occurs for a vertex in the majorization-center. This may be of interest in location analysis.

- There is a connection between the majorization-center and the notion of *Pareto location* discussed in the (location) literature, see e.g. [6]. To explain this connection define the special "distance" functions $f^k(v) = \sum_{j=1}^{k} d(v, \cdot)_{[j]}$ for each vertex v in the tree T and $k \leq n$. A Pareto location is a vertex v for which there is no other vertex (location) u such that $f^k(u) \leq f^k(v)$ for all $k = 1, 2, \ldots, n$, and with at least one strict such inequality. Then it follows from our results above that the set of Pareto locations in T is contained in the majorization-center. Note that the specific functions f^k defined here are, to our knowledge, not discussed in the location literature before. We refer to [6] for multicriteria location problems and to [3] for Pareto optimality in computational geometry (both papers also contain further references in these areas).

For further details on these questions we refer to [5].

3 Transportation matrices

Finally, we briefly explain how majorization may arise in connection with transportation matrices. The presentation is based on [4].

Let $R = (r_1, r_2, \ldots, r_m)$ and $S = (s_1, s_2, \ldots, s_n)$ be positive vectors of length m and n, respectively, such that $\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} s_j$. Let $\mathcal{N}(R, S)$ denote that set (class) of nonnegative matrices with row sum vector R and column sum vector S. These matrices are called *transportation matrices*, and they arise in the classical transportation problem in linear programming. The set $\mathcal{N}(R, S)$ is a polytope, called the *transportation polytope*. A matrix A in $\mathcal{N}(R, S)$ may be constructed using the following *North-West Corner rule* (see [8]). Let $a_{11} = \min\{r_1, s_1\}$; if $a_{11} = r_1$ (resp. $a_{11} = s_1$), let the remaining entries in the first row (resp. column) be zeros and update the first column sum (resp. row sum) by subtracting a_{11}. Delete the first row (resp. column) and proceed similarly with the smaller remaining matrix.

The following result gives a connection between transportation matrices and majorization.

Theorem 5 *Let R and S be nonincreasing, positive vectors of length n with $\sum_i r_i = \sum_j s_j$. Then the following are equivalent:*
(i) $S \preceq R$, *i.e.,* $\sum_{j=1}^{k} s_j \leq \sum_{j=1}^{k} r_j$ *$(1 \leq k \leq n-1)$.*
(ii) *There is an upper triangular matrix in $\mathcal{N}(R, S)$.*
(iii) *The North-West Corner rule for $\mathcal{N}(R, S)$ produces an upper triangular matrix.*

In the situation of this corollary, the matrix A produced by North-West Corner rule says how to redistribute the "mass" represented by the components of R in order to obtain S. This is done by moving mass to components with higher index; see the following example.

Example. If $R = (9, 2, 2, 1, 1)$ and $S = (5, 3, 3, 2, 2)$ the following upper triangular matrix is obtained by the North-West Corner rule

$$A = \begin{bmatrix} 5 & 3 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

So $S \preceq R$ and A indicates how to obtain S from R by moving 3 and 1 from the first component to the second and third components, respectively, etc.

□

We conclude with a remark:

- The theorem above allows a generalization. Let again a row sum vector R and column sum vector S be given. One may study the set of transportation matrices whose nonzero entries lie in a given *staircase pattern* W. Whether such a class is nonempty depends on the combination of R, S, and the given pattern. One may show that this class is nonempty if and only a certain set of majorization-like conditions hold for the vectors R and S, where the structure of these inequalities is determined by the given pattern W.

For further results and details, one may consult [4].

Acknowledgment. The author wishes to thank two referees for several useful comments to the paper.

References

[1] R. Bathia, *Matrix Analysis*, Graduate Texts in Mathematics vol. 169, Springer-Verlag, 1996.

[2] R.A. Brualdi. *Combinatorial Matrix Classes*. Cambridge University Press, 2006.

[3] V. Chepoi, K. Nouioua, Pareto envelopes in $\mathbb{R}^3$ under ℓ_1 and ℓ_∞ distance functions, Symposium on Computational Geometry SoCG'2007, pp. 284-293.

[4] G. Dahl, Transportation matrices with staircase patterns and majorization, Linear Algebra and its Appl., Volume 429 (Issue 7) (2008) pp. 1840-1850.

[5] G. Dahl, Majorization and distances in trees, Networks, Vol. 50 (Issue 4) (2007) pp. 251–257.

[6] M. Ehrgott, H.W. Hamacher and S. Nickel, Geometric methods to solve max-ordering location problems, Discrete Applied Mathematics 93 (1999) pp. 3–20.

[7] E. Erkut, Inequality measures for location problems, Location Sci. 1 (1993) pp. 199–217.

[8] D.G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, Massachusetts, second edition, 1989.

[9] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.