An algorithm for computing the matching capacity of a tree

Suren Arzumanyan

Abstract

For any bipartite graph $G = (\mathcal{X},\ \mathcal{Y},\ \mathcal{E})$ the matching capacity of G is

$$\gamma(G) = \lim_{n\to\infty} \frac{1}{n} \log \nu(G^{\otimes n}),$$

where $G^{\otimes n}$ is the n-th power of G and $\nu(G^{\otimes n})$ is its matching number. Ahlswede and Cai [1] showed that

$$\gamma(G) = \max \min_{(P,Q)\in\mathcal{K}(G)} \{H(P),\ H(Q)\},$$

where H is the entropy function and $\mathcal{K}(G)$ is the set of König-Hall pairs of distributions. In this paper an iterative method of computing the matching capacity of an arbitrary tree is presented. The following estimation is obtained

$$|H(Q_{t+1}) - \gamma(G)| \leq \frac{ct}{2^t},$$

where c is a constant and t is the number of iterations.

1 Basic concepts and auxiliary results

Let $G = (\mathcal{X},\ \mathcal{Y},\ \mathcal{E})$ be a tree, where $(\mathcal{X},\ \mathcal{Y})$ is the bipartition of G and $\mathcal{E}$ is the set of edges. Suppose $|\mathcal{X}| \leq |\mathcal{Y}|$. We use the following notations

$$\Gamma_G(v) = \{v' : (v, v') \in \mathcal{E}\}$$

$$\Gamma_G(S) = \bigcup_{v \in S} \Gamma_G(v)$$

$$\mathcal{K}(G) = \ \{(P, Q) \ : \ P \in \mathcal{P}(\mathcal{X}), \ Q \in \mathcal{P}(\mathcal{Y}), \ P(S) \ \leq \ Q(\Gamma_G(S)) \ \forall S \subset \mathcal{X}\}.$$

where $\mathcal{P}(\mathcal{X})$ (resp. $\mathcal{P}(\mathcal{Y})$) is the set of probability distributions on $\mathcal{X}$ (resp. $\mathcal{Y}$), and

$$P(S) = \sum_{x \in S} P(x),$$

$$Q(\Gamma_G(S)) = \sum_{y \in \Gamma_G(S)} Q(y).$$

The matching capacity $\gamma(G)$ of a bipartite graph $G = (\mathcal{X}, \ \mathcal{Y}, \ \mathcal{E})$ is given by

$$\gamma(G) = \max \min_{(P,Q) \in \mathcal{K}(G)} \{H(P), \ H(Q)\},$$

where

$$H(P) = -\sum_{x \in X} P(x) \log P(x)$$

is the well known entropy function.

Observation 1 *Let* $P_u \in \mathcal{P}(\mathcal{X})$ *be the uniform distribution. If there exists* $Q \in \mathcal{P}(\mathcal{Y})$ *such that*

$$(P_u, Q) \in \mathcal{K}(G) \text{ and } H(P_u) \leq H(Q)$$

then

$$\gamma(G) = H(P_u) = \log |\mathcal{X}|.$$

Proof. For every $P \in \mathcal{P}(\mathcal{X})$

$$0 \leq H(P) \leq \log |\mathcal{X}| = H(P_u). \ \blacksquare$$

Given two distributions

$$P = (P(0), \ldots, P(\alpha - 1))$$

and

$$Q = (Q(0), \ldots, Q(\alpha - 1)),$$

we say that P majorizes Q and write $P \succ Q$, if

$$\sum_{i=0}^{k} P[i] \geq \sum_{i=0}^{k} Q[i] \quad \text{for} \quad k = 0, 1, \ldots, \alpha - 1$$

where $P[i]$(resp. $Q[i]$) is the i-th largest component of P(resp. Q).

A function $\varphi : \mathcal{P}(\mathcal{X}) \longrightarrow \mathbf{R}$ is Schur-concave, if $P \succ Q$ implies

$$\varphi(P) \leq \varphi(Q),$$

and it is strictly Schur-concave, if $P \succ Q$ and $P \neq Q$ imply

$$\varphi(P) < \varphi(Q).$$

It is well known [2]

Proposition 1 *The entropy function H is strictly Schur-concave.* ■

Construction 1 *Given the uniform probability distribution $P \in \mathcal{P}(\mathcal{X})$ construct a probability distribution $Q \in \mathcal{P}(\mathcal{Y})$ such that*

$$(P, Q) \in \mathcal{K}(G)$$

and

$$\max_{Q' :\ (P,Q') \in \mathcal{K}(G)} H(Q') = H(Q). \quad ■$$

This is a linear programming problem. Given linear constraints described by conditions $P(S) \leq Q(\Gamma_G(S))$ for all $S \subset \mathcal{X}$ of the set $\mathcal{K}(G)$, we want to maximize entropy H. But we have exponentially many constraints since we have $2^{|\mathcal{X}|}$ many subsets. Therefore the known methods are not effective and we give the following construction.

Remark 1 *A graph with the bipartition $(x, \Gamma_G(x))$ for $x \in X$ we will call a star.*

Definition 1 *The local root of a star is the closest to the current vertex vertex of the star.*

Definition 2 *The branch corresponding to a vertex x is the subtree $T = (V(T), E(T))$ with the following properties*

1. *$x \in V(T)$*
2. *any star of T has a positive value in its local root*
3. *for any star $S = (V(S), E(S))$ of T with $x \in V(S)$, x is the local root of S*
4. *any star of T contains a vertex with the value equal to the value of x*
5. *T is the maximal tree with properties 1-4, i.e. any subtree T' satisfying 1-4 is a subtree of T.*

For a vertex $x \in X$ consider the following sequence of values

$$Q(1) = Q(2) = \ldots = Q(i_1) <$$

$$Q(i_1 + 1) = \ldots = Q(i_1 + i_2) < \ldots \tag{1}$$

$$Q(i_1 + i_2 + \ldots + i_{n-1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{n-1} + i_n)$$

where

$$Q(1) = Q(2) = \ldots = Q(i_1)$$

are the smallest values of the incident vertices $\Gamma_G(x)$ of the vertex x and

$$Q(i_1 + i_2 + \ldots + i_{k-1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{k-1} + i_k)$$

are the smallest values greater than $Q(i_{k-1})$ on the union of the branches corresponding to the vertices associated with the previous values

$$Q(1), \ldots, Q(i_1 + i_2 + \ldots + i_{k-1})$$

(If $Q(1) = Q(2) = \ldots = Q(i_1)$ are zeros, then $Q(i_1 + 1) = \ldots = Q(i_1 + i_2)$ are the smallest nonzero values of $\Gamma_G(x)$).

For $k = 1, \ldots, n$, set

$$\omega(k) = \frac{\frac{1}{|X|} + i_1 Q(i_1) + i_2 Q(i_1 + i_2) + \ldots + i_k Q(i_1 + i_2 + \ldots + i_k)}{i_1 + i_2 + \ldots + i_k}.$$

Definition 3 *The union of the branches associated with sequence (1) is called the current tree.*

Definition 4 $\omega(k)$ *is called the average value and the vertices associated with the values*

$$Q(1), \ldots, Q(i_1 + i_2 + \ldots + i_k)$$

the averaging vertices of the vertex x, where

$$k = \begin{cases} n & \text{if } \omega(l) > Q(i_1 + \ldots + i_{l+1}) \\ & \text{for } l = 1, \ldots, n-1 \\ \\ \min \{ l \ : \ \omega(l) \leq Q(i_1 + \ldots + i_{l+1}) \} & \text{otherwise} \end{cases} .$$

For a star with the local root y consider the following subsequence of sequence (1):

$$\begin{gathered} Q(i_1 + i_2 + \ldots + i_{j_1 - 1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{j_1 - 1} + i_{j_1}) < \ldots \\ Q(i_1 + i_2 + \ldots + i_{j_2 - 1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{j_2 - 1} + i_{j_2}) < \ldots \\ Q(i_1 + i_2 + \ldots + i_{j_n - 1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{j_n - 1} + i_{j_n}) \end{gathered} \quad (2)$$

where

$$Q(i_1 + i_2 + \ldots + i_{j_1 - 1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{j_1 - 1} + i_{j_1})$$

are the values of the vertices from the branch of y having the same value as y and

$$Q(i_1 + i_2 + \ldots + i_{j_k - 1} + 1) = \ldots = Q(i_1 + i_2 + \ldots + i_{j_k - 1} + i_{j_k})$$

are the values from (1) on the union of branches corresponding to the vertices associated with the previous values

$$Q(i_1 + i_2 + \ldots + i_{j_1 - 1} + 1), \ldots, Q(i_1 + i_2 + \ldots + i_{j_k - 1}).$$

Definition 5 *The union of the branches associated with sequence (2) of a star is called the tree associated with this star.*

Definition 6 *The weight of a star is the value*

$$w = \sum_{t=1}^{k} (\omega(k) - Q(i_1 + i_2 + \ldots + i_{j_t - 1} + i_{j_t})) i_{j_t},$$

where $\omega(k)$ is the average value of the current vertex.

Definition 7 *A star is correct if its weight is not greater than the value of this star contained in the local root.*

Step i.

If there are vertices not chosen before in **Step i**, then take a vertex $x \in X$ with the minimal degree from them and call it the current vertex.

Step ii.

If all vertices are chosen, then stop.

Step iii.

If there is an incorrect star in the current tree, then consider a new current tree obtained by cutting from the current tree all smallest trees associated with the incorrect stars (smallest means that there are no incorrect stars in these trees).

Step iv.

If there are no incorrect stars, then change the values of the averaging vertices to $\omega(k)$ and go to **Step vi**.

Step v.

Let γ be the sum of all values which have these incorrect stars in the local roots. Let $\omega(k)$ be calculated with respect to the new current tree (we do not consider the cut trees) taking instead of $\frac{1}{|X|}$ the value $\frac{1}{|X|} - \gamma$ and go to **Step iii**.

Step vi.

If there is a cut incorrect star obtained by **Step iii**, then consider the closest to the current vertex cut incorrect star and the tree associated with it (closest means there is no incorrect star, which was not considered in **Step vi** before, between this incorrect star and the current vertex). Let $\omega(k)$ be calculated with respect to this tree taking instead of $\frac{1}{|X|}$ the value γ_i of the incorrect star in the local root. Consider this tree as the current tree and go to **Step iii** taking instead of $\frac{1}{|X|}$ the value γ_i.

Step vii.

If there are no cut incorrect stars, then go to **Step i**.

Remark 2 *Going from **Step vi** to **Step iii**, the correctness of the incorrect stars should be checked at first. If an incorrect star becomes correct for these new parameters, then all correct stars of the tree associated with this incorrect star will remain correct.*

Construction 2 *Given a probability distribution $Q \in \mathcal{P}(\mathcal{Y})$ construct the probability distributions $Q_1 \in \mathcal{P}(\mathcal{Y})$ and $P_1 \in \mathcal{P}(\mathcal{X})$ such that*

$$\sum_{y \in \mathcal{Y}} |Q(y) - Q_1(y)| = \alpha,$$

$$\max_{Q' \,:\sum_{y \in \mathcal{Y}} |Q(y)-Q'(y)|=\alpha} H(Q') = H(Q_1)$$

where

$$\alpha \leq \frac{1}{2}\sum_{y\in\mathcal{Y}}|Q(y)-\frac{1}{|\mathcal{Y}|}|,$$

$$(P_1,\ Q_1)\in\mathcal{K}(G)$$

and

$$\max_{P':\ (P',Q_1)\in\mathcal{K}(G)} H(P') = H(P_1)\ \blacksquare.$$

Suppose Q is distributed in the following way

$$Q[1]=Q[2]=\ldots=Q[i_1]>$$

$$Q[i_1+1]=Q[i_1+2]=\ldots=Q[i_1+i_2]>$$

$$Q[i_1+i_2+1]=\ldots$$

We use the following notation

$$\delta_k=(Q[i_1]-Q[i_1+1])i_1+$$

$$(Q[i_1+i_2]-Q[i_1+i_2+1])(i_1+i_2)+\ldots$$

$$(Q[i_1+\ldots+i_k]-Q[i_1+\ldots+i_k+1])(i_1+\ldots+i_k)$$

and $\delta_0=0$. If

$$\delta_k<\frac{\alpha}{2}\leq\delta_{k+1}$$

then change the values

$$Q[1],\ldots,Q[i_1+i_2+\ldots+i_{k+1}]$$

to

$$Q'[1]=\ldots=Q'[i_1+i_2+\ldots+i_{k+1}]=Q[i_1+\ldots+i_{k+1}]-\frac{\alpha/2-\delta_k}{i_1+\ldots+i_{k+1}}.$$

We used this procedure to push the largest values of Q on value $\frac{\alpha}{2}$ down. Now we use the similar procedure to push the smallest values of Q on value $\frac{\alpha}{2}$ up.

Proceeding in this way we shall obtain the desired probability distribution Q_1. The property

$$\sum_{y \in \mathcal{Y}} |Q(y) - Q_1(y)| = \alpha$$

is satisfied by the construction of Q_1. The second property

$$\max_{Q'\colon\ \sum_{y\in\mathcal{Y}} |Q(y)-Q'(y)|=\alpha} H(Q') = H(Q_1)$$

is satisfied since the entropy function H is $Schur - concave$.

To construct P_1 we use the following notations

$$S_1 := \bigcup_{t=1}^{i_1} q(t)$$

where $q(1), ..., q(i_1)$ are vertices associated with values $Q[1], ..., Q[i_1]$ and for $l = 2, ..., k+1$

$$S_l := \bigcup_{j=1}^{i_l} q(i_1 + \ldots + i_{l-1} + j)$$

where $q(i_1 + \ldots + i_{l-1} + j),\ j = 1, ..., i_l$ are the vertices associated with values $Q[i_1 + \ldots + i_{l-1} + j],\ j = 1, ..., i_l$. Subtract from the values associated with vertices $\Gamma_G(S_1)$ the value

$$\frac{(Q[1] - Q'[1])i_1}{|\Gamma_G(S_1)|}.$$

Subtract from the values associated with vertices $\Gamma_G(S_l) - \Gamma_G(S_{l-1})$ the value

$$\frac{(Q[i_1 + \ldots + i_l] - Q'[1])i_l}{|\Gamma_G(S_l) - \Gamma_G(S_{l-1})|}.$$

By this procedure we decrease the values of the vertices associated with the values pushed down during the construction of Q_1. Now we use the similar procedure to increase the values of the vertices associated with the values pushed up during the construction of Q_1. Proceeding in this way we shall obtain the desired probability distribution P_1. The property

$$(P_1, Q_1) \in \mathcal{K}(G)$$

is satisfied since the values of S_l are the same and the values of $\Gamma_G(S_l) - \Gamma_G(S_{l-1})$ are the same. Therefore to keep the property of $\mathcal{K}(G)$ it is enough to apply the procedure described above. The second property

$$\max_{P':\ (P',Q_1)\in\mathcal{K}(G)} H(P') = H(P_1).$$

is satisfied since the entropy function is $Schur-concave$.

Remark 3 $H(Q) < H(Q_1)$ *since the entropy function* H *is strictly* $Schur-concave$.

2 Iterative procedure and convergence

For the uniform probability distribution $P_1 \in \mathcal{P}(\mathcal{X})$ using $Construction$ 1 construct a probability distribution $Q_1 \in \mathcal{P}(\mathcal{Y})$ such that

$$(P_1, Q_1) \in \mathcal{K}(G)$$

and

$$\max_{Q':\ (P_1,Q')\in\mathcal{K}(G)} H(Q') = H(Q_1).$$

If

$$H(P_1) \le H(Q_1),$$

then by $Observation$ 1

$$\gamma(G) = H(P_1).$$

If

$$H(P_1) > H(Q_1),$$

then set $t = 1$ and

$$\alpha = \frac{1}{2}\sum_{y\in\mathcal{Y}}\left|Q_1(y) - \frac{1}{|\mathcal{Y}|}\right|.$$

Step i.

Using *Construction* 2 construct probability distributions $Q_{t+1} \in \mathcal{P}(\mathcal{Y})$ and $P_{t+1} \in \mathcal{P}(\mathcal{X})$ such that

$$\sum_{y \in \mathcal{Y}} |Q_t(y) - Q_{t+1}(y)| = \alpha,$$

$$\max_{Q' \colon \sum_{y \in \mathcal{Y}} |Q_t(y) - Q'(y)| = \alpha} H(Q') = H(Q_{t+1}),$$

$$(P_{t+1}, Q_{t+1}) \in \mathcal{K}(G)$$

and

$$\max_{P' \colon (P', Q_{t+1}) \in \mathcal{K}(G)} H(P') = H(P_{t+1})$$

Step ii.

If

$$H(P_{t+1}) > H(Q_{t+1}),$$

then set $t := t + 1$, $\alpha := \frac{\alpha}{2}$ and go to Step i.

Step iii.

If

$$H(P_{t+1}) < H(Q_{t+1}),$$

then set $\alpha := \frac{\alpha}{2}$ and go to Step i.

Step iv.

If

$$H(P_{t+1}) = H(Q_{t+1}),$$

then $\gamma(G) = H(P_{t+1})$.

Remark 4 *By the definition of* $\mathcal{K}(G)$ *and Construction* 2 *for every* $Q \in \mathcal{P}(\mathcal{Y})$ *such that* $Q \prec Q_t$ *the following inequality*

$$H(P) < H(P_t)$$

holds, where P *is the probability distribution constructed by Construction* 2 *for given* Q *such that*

$$(P,\ Q) \in \mathcal{K}(G)$$

and

$$\max_{P':\ (P',Q)\in\mathcal{K}(G)} H(P') = H(P). \blacksquare$$

Theorem 1 *If* ***Step iv*** *never happens, then the sequences* $H(Q_t)$ *and* $H(P_t)$ *produced by the iterative procedure converge to* $\gamma(G)$ *and the following*

$$|H(Q_{t+1}) - \gamma(G)| \leq \frac{c\ t}{2^t} \tag{3}$$

holds, where c *is a constant .*

Proof. Set

$$Q_{op.} := \lim_{t\to\infty} Q_t \ \text{ and } \ P_{op.} := \lim_{t\to\infty} P_t.$$

By the iterative procedure and *Remarks* 3 and 4 we have

$$H(Q_{op.}) = H(P_{op.}).$$

By *Remark* 4 for any $Q \in \mathcal{P}(\mathcal{Y})$ such that $Q \prec Q_{op.}$ we have

$$H(P) < H(P_{op.})$$

where

$$H(P) = \max_{P':\ (P',Q)\in\mathcal{K}(G)} H(P')$$

and therefore

$$H(Q_{op.}) = \gamma(G).$$

To prove (3) we use the following lemma [3]:

Lemma 1 *If P and Q are two distributions on $\mathcal{Y}$ such that*

$$\sum_{y\in\mathcal{Y}} |P(y) - Q(y)| = \theta \leq \frac{1}{2}$$

then

$$|H(P) - H(Q)| \leq -\theta \log \frac{\theta}{|\mathcal{Y}|}. \quad \blacksquare$$

Using this lemma and the facts that

$$\sum_{y\in\mathcal{Y}} |Q_t(y) - Q_{t+1}(y)| = \frac{1}{2^t} \sum_{y\in\mathcal{Y}} |Q_1(y) - \frac{1}{|\mathcal{Y}|}|$$

and

$$\sum_{y\in\mathcal{Y}} |Q_{t+1}(y) - Q_{op.}(y)| < \sum_{y\in\mathcal{Y}} |Q_{t+1}(y) - Q_t(y)|$$

we obtain (3). $\blacksquare$

The proof of **Step iv**. is the same.

Remark 5 *For bipartite graphs the statement of Remark 4 is in general not satisfied and therefore the convergence in this case can not be shown.*

References

[1] R. Ahlswede and N. Cai, Information and control: matching channels, IEEE Trans. Inform. Theory 43, No. 1, 48-70, 1997.

[2] A.W. Marshall and I. Olkin, Inequalities: theory of majorization and its applications, Vol. 143 in Mathematics in Science and Engineering, Academic 1979.

[3] I. Csiszár and J. Körner, Information theory: coding theorems for discrete memoryless systems, Academic Press, 1981.