Stable matchings and linear programming

Rakesh V. Vohra*

Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208

This article provides a brief summary of the well-known stable matching problem from the point of view of linear programming. One contribution is an interpretation of the Gale–Shapley proposal algorithm as a dual ascent algorithm for an appropriate linear programming formulation of the stable matching problem.

Keywords: Blocking pair, convex hull, linear programming, stable matching.

Introduction

THE stable matching problem was introduced by Gale and Shapley[1] as a model of how to assign students to colleges. It is a classic that has inspired a flood of papers exploring generalizations, variations and applications. For a comprehensive account of these, see Roth and Sotomayor[2] as well as Roth[3].

A stable matching is a matching in a bipartite graph that satisfies additional conditions. Just as we have a linear inequality description of the convex hull of all matchings in a bipartite graph, it is natural to ask if such a description is possible for the convex hull of stable matchings. Vande Vate[4] provided one. His proof (as do subsequent ones) assumes the existence of a stable matching. Existence was established by Gale and Shapley[1] via their proposal algorithm. What has nagged me is that one should be able to obtain existence of a stable matching directly from Vande Vate's characterization of the convex hull of stable matchings. This article does just that. I associate a linear program with the linear inequality description of the convex hull of stable matchings. I then show that an appropriate dual ascent algorithm for this linear program produces a stable matching. Indeed, the dual ascent algorithm resembles the Gale–Shapley proposal algorithm; which I find satisfying.

I begin by introducing notation and stating the stable matching problem. Subsequently, I review the proposal algorithm and Vand Vate's characterization. This will make the article self-contained.

Stable matching problem

Given is a set M of men and a set W of women. Each $m \in M$ has a strict preference ordering over the women in W and each $w \in W$ has a strict preference ordering over the men in M. The preference ordering of agent i is denoted $\succ_i$ and $x \succ_i y$ means agent i ranks x above y. One can accommodate the possibility of an agent choosing to remain single by including for each man (woman) a dummy woman (man) in the set $W(M)$ that corresponds to being single (or matched with oneself). The dummy woman associated with man m will rank m first and all other men below him. All men other than m will rank this dummy woman (in any order) below their corresponding dummy woman. With this construction we can assume $|M| = |W|$.

A matching is an assignment of men to women such that each man is assigned or paired with exactly one woman and no woman is assigned to more than one man. Denote a matching by μ. The woman matched to man m in the matching μ is denoted $\mu(m)$. Similarly, $\mu(w)$ is the man matched to woman w.

A matching μ is said to be blocked by the pair (m, w') if

1. $\mu(m) = w$,
2. $\mu(m') = w'$,
3. And $w' \succ_m w$ and $m \succ_{w'} m'$.

The pair (m, w') is called a *blocking pair*. A matching that has no blocking pairs is called *stable*.

Example 1. *The preference orderings of men and women are shown in the table below.*

$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$	$\succ_{w_1}$	$\succ_{w_2}$	$\succ_{w_3}$
w_2	w_1	w_1	m_1	m_3	m_1
w_1	w_3	w_2	m_3	m_1	m_3
w_3	w_2	w_3	m_2	m_2	m_2

Consider the matching $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$. *It is not stable because* (m_1, w_2) *is a blocking pair. The matching* $\{(m_1, w_1), (m_3, w_2), (m_2, w_3)\}$, *however, is stable.*

The presence of dummy men and women means that no stable matching (if it exists) forces an agent into a match with a partner whom he/she ranks below being single. That agent and their corresponding dummy on the other side would form a blocking pair.

The proposal algorithm

It is not at all obvious that a stable matching exists. The main result of Gale and Shapley[1] is that a stable matching

*e-mail: r-vohra@kellogg.northwestern.edu

always exists. The proof is constructive via an elegant algorithm called the deferred acceptance algorithm.

Deferred acceptance algorithm (male-propose version)

1. First, each man proposes to his top-ranked choice.
2. Next, each woman who has received at least two proposals keeps (tentatively) her top-ranked proposal and rejects the rest.
3. Then, each man who has been rejected, proposes to his top-ranked choice amongst the women who have not rejected him.
4. Again each woman who has at least two proposals (including the ones kept from previous rounds), keeps her top-ranked proposal and rejects the rest.
5. The process repeats until no man has a woman to propose to, or each woman has at most one proposal.

Theorem 1. *The male-propose deferred acceptance algorithm terminates in a stable matching.*

Proof. When it terminates, it clearly does so in a matching. Suppose the matching is blocked by the pair (m_1, w_1) with m_1 matched to w_2, say, and w_1 matched to m_2. Since (m_1, w_1) is blocking and $w_1 \succ_{m_1} w_2$, in the proposal algorithm, m_1 would have proposed to w_1 before w_2. Also, m_1 was not matched with w_1 by the algorithm, because w_1 received a proposal from a man that she ranked higher than m_1. As the algorithm matches her to m_2 it follows that $m_2 \succ_{w_1} m_1$. This contradicts the fact that (m_1, w_1) is a blocking pair. □

One could describe a deferred acceptance algorithm where the females propose and the outcome would also be a stable matching possibly different from the one returned using the male-propose deferred acceptance algorithm. Thus, not only is a stable matching guaranteed to exist, but there can be more than one. If there can be more than one stable matching, is there a reason to prefer one to another?

A matching μ is *male-optimal* if there is no stable matching ν such that $\nu(m) \succ_m \mu(m)$ or $\nu(m) = \mu(m)$ for all m with $\nu(j) \succ_j \mu(j)$ for at least one $j \in M$. Similarly define *female-optimal*.

Theorem 2. *The stable matching produced by the male-propose deferred acceptance algorithm is male-optimal.*

Proof. Let μ be the matching returned by the male-propose deferred acceptance algorithm. Suppose μ is not male optimal. Then, there is a stable matching ν such that $\nu(m) \succ_m \mu(m)$ or $\nu(m) = \mu(m)$ for all m with $\nu(j) \succ_j \mu(j)$ for at least one $j \in M$. Therefore, in the application of the proposal algorithm there must be an iteration where some man j proposes to $\nu(j)$ before $\mu(j)$ since $\nu(j) \succ_j \mu(j)$ and is rejected by woman $\nu(j)$. Consider the first such iteration. Since woman $\nu(j)$ rejects j, she must have received a proposal from a man i she prefers to man j. Since this is the first iteration at which a male is rejected by his partner under ν, it follows that man i ranks woman $\nu(j)$ higher than $\nu(i)$. Summarizing, $i \succ_{\nu(j)} j$ and $\nu(j) \succ_i \nu(i)$, implying that ν is not stable, a contradiction. □

You can replace the word 'male' by the word 'female' in the statement of the theorem. There does not always exist a stable matching that is simultaneously optimal for the males and females.

The convex hull of stable matchings

Vande Vate[4] identified a collection of linear inequalities that described the convex hull of stable matchings. I will describe these inequalities and show using an elegant rounding argument of Teo and Sethurman[5] that they describe the convex hull of stable matchings.

For each man m and woman w, let $x_{mw} = 1$, if man m is matched with woman w and 0 otherwise. Then, every stable matching must satisfy the following

$$\sum_{w \in W} x_{mw} = 1 \quad \forall m \in M, \tag{1}$$

$$\sum_{m \in M} x_{mw} = 1 \quad \forall w \in W, \tag{2}$$

$$\sum_{j \prec_m w} x_{mj} + \sum_{j \prec_w m} x_{iw} + x_{mw} \le 1 \quad \forall m \in M, w \in W, \tag{3}$$

$$x_{mw} \ge 0 \quad \forall m \in M, w \in W. \tag{4}$$

Let P be the polyhedron defined by eqs (1)–(4).

Constraints (1) and (2) ensure that each agent is matched with exactly one other agent of the opposite sex. Constraint (3) ensures stability by ruling out blocking pairs (call it the blocking constraint). To see why, suppose $\sum_{j \prec_m w} x_{mj} = 1$ and $\sum_{i \prec_w m} x_{iw} = 1$. Then, man m is matched to a woman j that he ranks below w. Similarly, woman w is matched to a man she ranks below m. This would make the pair (m, w) a blocking pair.

The following lemma is from Roth *et al.*[6].

Lemma 1. *Suppose* $P \ne \emptyset$. *Let* $x \in P$. *Then,* $x_{mw} > 0$ *implies that*

$$\sum_{j \prec_m w} x_{mj} + \sum_{j \prec_w m} x_{iw} + x_{mw} = 1.$$

Proof. Consider $\min\{\sum_i \sum_j x_{ij} : x \in P\}$. The dual to this program is

$$\max \sum_{i \in M} \alpha_i + \sum_{j \in W} \beta_j - \sum_{i \in M} \sum_{j \in W} \nu_{ij}$$

subject to

$$\alpha_i + \beta_j - \sum_{k \in W : k \succeq_{ij}} \nu_{ik} - \sum_{k \in M : k \succ_j i} \nu_{kj} \le 1 \quad \forall i \in M,\ j \in W,$$

$$\nu_{ij} \ge 0 \quad \forall i \in M,\ j \in W.$$

Set $\alpha_i = \sum_{j \in W} \nu_{ij}$ and $\beta_j = \sum_{i \in M} \nu_{ij}$. Substituting this into the dual constraints yields

$$\sum_{k \in W : k \prec_i j} \nu_{ik} + \sum_{k \in M : k \prec_j i} \nu_{kj} + \nu_{ij} \le 1 \quad \forall i \in M,\ j \in W,$$

Choose any $x^* \in P$ and set $\nu_{ij} = x^*_{ij}$. This choice of ν is clearly dual feasible. It has an objective function value of $\sum_{i \in M} \sum_{j \in W} x^*_{ij}$ and so is dual optimal. The lemma now follows by complementary slackness. □

The proof below is due to Teo and Sethuraman[5].

Theorem 3. *Suppose $P \neq \emptyset$. Then, P is the convex hull of stable matchings.*

Proof. Choose any weight vector $\{c_{ij}\}_{i \in M, j \in W}$ and let

$$x^* \in \arg\max \left\{ \sum_i \sum_j c_{ij} x_{ij} : x \in P \right\}.$$

With each member of $M \cup W$ we associate an interval (0, 1]. For each $i \in M$, partition the associated interval $(0, 1]_i$ into subintervals of length x^*_{ij} for all $j \in W$. Arrange these subintervals left to right by man i's decreasing preference over W. For each woman $j \in W$, partition the associated interval $(0, 1]_j$ into subintervals of length x^*_{ij} for all $i \in M$. Arrange these subintervals left to right by woman j's increasing preference over M.

Lemma 1 means the subinterval spanned by x^*_{ij} in $(0, 1]_i$ and $(0, 1]_j$ coincides. Pick a random number U uniformly in (0, 1] and construct a matching in the following way:

1. Match $i \in M$ to $k \in W$, if U lies in the subinterval of $(0, 1]_i$ spanned by x^*_{ik}.
2. Match $j \in W$ to $i \in M$, if U lies in the subinterval of $(0, 1]_j$ spanned by x^*_{ij}.

By Lemma 1, $i \in M$ is matched to $j \in W$, iff $j \in W$ is matched to $i \in M$. Furthermore, no two men can be matched to the same woman, and similarly, no two women can be matched to the same man. So the above procedure does return a feasible matching.

To show this matching is stable, consider man i who is matched to $j \in W$, but prefers $k \in W$. Then, in $(0, 1]_i$, the subinterval corresponding to x^*_{ik} is to the left of the subinterval corresponding to x^*_{ij}. Because $i \in M$ is not matched to $k \in W$, it means that U is to the right of the subinterval corresponding to x^*_{ik} in $(0, 1]_i$. Therefore, U is to the left of the subinterval corresponding to x^*_{ik} in $(0, 1]_j$. In other words, $j \in W$ is matched to someone she prefers to $i \in M$.

Set $X^U_{ij} = 1$, iff man i is matched to woman j by the randomized scheme above. Then

$$E\left(\sum_{i \in M} \sum_{j \in W} c_{ij} X^U_{ij} \right) = \sum_{i \in M} \sum_{j \in W} c_{ij} E(X^U_{ij}) = \sum_{i \in M} \sum_{j \in W} c_{ij} x^*_{ij}.$$

Thus, we have exhibited a probability distribution over stable matchings whose expected objective function value coincides with $c \cdot x^*$. It follows then, that there is a stable matching with objective function value $c \cdot x^*$. □

A dual ascent algorithm

Theorem 2 assumes that $P \neq \emptyset$. We know that $P \neq \emptyset$ by virtue of the proposal algorithm. It would be nice to verify $P \neq \emptyset$ directly from eqs (1)–(4). I do this below.

If woman j is man i's first choice, set $r_{ij} = n$. If woman j is man i's second choice, set $r_{ij} = n - 1$ and so on. Consider

$$\max \left\{ \sum_{m \in M} \sum_{w \in M} r_{mw} x_{mw} : x \in P \right\}.$$

Let ν be the dual multiplier associated with the blocking constraints and p the dual multiplier associated with the constraint $\sum_{w \in W} x_{mw} = 1$ for all m. Consider the following Lagrangean relaxation of the optimization problem

$$\max \sum_{m \in M} \sum_{w \in W} \left[r_{mw} - p_w - \nu_{mw} - \sum_{k \in W : k \succ_m w} \nu_{mk} - \sum_{k \in M : k \succ_w m} \nu_{km} \right] x_{mw}$$

$$\text{s.t.} \sum_{w \in W} x_{mw} = 1 \ \forall m \in M, \quad x_{mw} \ge 0 \ \forall (m, w).$$

This relaxation is easy to solve. For each man $m \in M$, choose the woman $w \in W$ that maximizes

$$r_{mw} - p_w - \nu_{mw} - \sum_{k \in W : k \succ_m w} \nu_{mk} - \sum_{k \in M : k \succ_w m} \nu_{km}.$$

In case of a tie, choose the top-ranked woman.

Let (p^t, ν^t) be the value of the multipliers at the start of iteration t. Set $(p^0, \nu^0) = 0$. Denote by x^t the optimal choice of x given the multipliers (p^t, ν^t). If $x^t_{mw} = 1$, we will say that man m proposed to woman w. Given x^t, we adjust (p^t, ν^t) by $(\hat{p}, \hat{\nu})$ as follows.

1. $\hat{p} = 0$ throughout.
2. At iteration t, let $\hat{v}_{ij} = x^t_{ij}$ for all i, j.
3. Set $(p^{t+1}, v^{t+1}) = (p^t, v^t) + (\hat{p}, \hat{v}) = (p^t, v^t) + (0, x^t_{ij})$.
4. Set $r^{t+1}_{mw} = r^t_{mw} - v^t_{mw} - \sum_{k \in W: k \succ_m w} v^t_{mk} - \sum_{k \in M: k \succ_w m} v^t_{km}$.

Notice that we can rewrite this to read

$$r^{t+1}_{mw} = r^t_{mw} - x^t_{mw} - \sum_{k \in W: k \succ_m w} x^t_{mk} - \sum_{k \in M: k \succ_w m} x^t_{km}.$$

Stop once x^t is a matching. I will argue below that this procedure terminates in a matching. By complementary slackness this will be an optimal stable assignment.

Call r^t the dual adjusted rank. Suppose at iteration t, $x^t_{mw} = 1$. In the cases below I examine how r^{t+1} differs from r^t. The goal is to show that the dual adjusted rank declines in a way that each man will propose to the women in much the same order as they would under the male-propose deferred acceptance algorithm.

In the dual ascent algorithm, when man m proposes to woman w at iteration t, she must be his ranked choice (as measured by the dual adjusted rank) at that time. If man m is woman w's top choice among those currently proposing to her, then man m's dual adjusted rank for w (as well as all women he ranks below w) will go down by 1. However, the dual adjusted rank of all women that man m ranks above w will go down by more than 1. Thus, at the next iteration, woman w will still be man m's highest ranked woman and he continues to propose to her.

If woman w rejects man m, then the dual adjusted rank for woman w goes down by 2. However, the dual adjusted rank of women he ranks below w that do not have a proposal from a man they rank higher, will go down only by 1. Thus, at the next iteration he does not propose to his next most preferred woman. Rather, he 'skips' over her if she already has a proposal from a man she ranks higher than m.

Case 1: If among all $i \in M$ such that $x^t_{iw} = 1$, man m is woman w's top-ranked choice, $r^{t+1}_{mw} = r^t_{mw} - 1$.

Because i is woman w's top-ranked choice, it follows that $\sum_{k \in M: k \succ_w m} x^t_{km} = 0$. Because $x^t_{mw} = 1$, it follows that $\sum_{k \in W: k \succ_m w} x^t_{mk} = 0$.

Case 2: If among all $i \in M$ such that $x^t_{iw} = 1$, man m is not woman w's top-ranked choice, $r^{t+1}_{mw} \leq r^t_{mw} - 2$.

Because i is not woman w's top-ranked choice, it follows that $\sum_{k \in M: k \succ_w m} x^t_{km} \geq 1$. Because $x^t_{mw} = 1$, it follows that $\sum_{k \in W: k \succ_m w} x^t_{mk} = 0$.

Case 3: If $w \succ_m j$ and $x^t_{kj} = 0$ for all $k \in M$ such that $k \succ_j m$, then $r^{t+1}_{mj} = r^t_{mj} - 1$.

Clearly $\sum_{k \in M: k \succ_w m} x^t_{km} = 0$ and $x^t_{mj} = 0$. Also $\sum_{k \in W: k \succ_m w} x^t_{mk} = 1$.

Case 4: If $w \succ_m j$ and $x^t_{kj} = 1$ for at least one $k \in M$ such that $k \succ_j m$, then $r^{t+1}_{mj} \leq r^t_{mj} - 2$.

Clearly $\sum_{k \in M: k \succ_w m} x^t_{km} \geq 1$ and $x^t_{mj} = 0$. Also $\sum_{k \in W: k \succ_m w} x^t_{mk} = 1$.

Case 5: If $j \succ_m w$, then $r^{t+1}_{mj} = r^t_{mj} - \sum_{k \in M: k \succ_j m} x^t_{km}$. Clearly, $x^t_{mj} = 0$. Also $\sum_{k \in W: k \succ_m w} x^t_{mk} = 0$.

Let $P^t = \{w \in W: \sum_{m \in M} x_{mw} \geq 1\}$. I show first that $P^t \subseteq P^{t+1}$. Consider first the case $t = 0$. Suppose man m proposed to woman w and man m is woman w's top choice among those who have proposed. Then the dual adjusted rank of woman w by man m declines by exactly 1 (see case 1). The dual adjusted rank of all other women declines by at least 1 (cases 2–4). Note that case 5 does not apply in this iteration. Therefore, in the next iteration man m will continue to propose to woman w. More generally, any woman who receives a proposal in iteration 0 continues to receive a proposal in iteration 1. Suppose this holds until iteration t. Suppose man m proposes to woman w in iteration t and man m is woman w's top choice among those who have proposed at iteration t. By cases 1–4 the dual adjusted rank for all woman ranked below w declines by at least 1. Consider a woman $j \succ_m w$. By case 5,

$$r^{t+1}_{mj} = r^t_{mj} - \sum_{k \in M: k \succ_j m} x^t_{km}.$$

However $\sum_{k \in M: k \succ_j m} x^t_{kw} \geq 1$ because any woman who received a proposal before iteration t continues to receive a proposal at iteration t. Thus the dual adjusted rank of all women ranked above w declines by at least 1. Because man m did not proposes to these women in iteration t their dual adjusted rank must be at least 2 smaller than r^t_{mw}. Thus, woman w continues to have the highest dual adjusted rank for man m and at iteration $t + 1$, man m will continue to propose to woman w.

Next, I show that there must be some t such that $P^t = W$. If not, then, it must be the case that $P^t = P^{t+1} = \ldots \subset W$. Because $P^t \subset W$, there must be a woman w who is proposed to by more than one man. Consider a man m who has proposed to her at iteration t and who is not her top choice among those proposing to her. Notice the following:

1. By case 2, $r^{t+1}_{mw} \leq r^t_{mw} - 2$.
2. By case 4, for all $j \in P^t$ such that $w \succ_m j$, $r^{t+1}_{mj} \leq r^t_{mj} - 2$.
3. By case 5, for all $j \in P^t$ such that $j \succ_m w$, $r^{t+1}_{mj} \leq r^t_{mj} - 1$.
4. By case 3, for all $j \notin P^t$, $r^{t+1}_{mj} = r^t_{mj} - 1$. Note, for all such j it must be that $w \succ_m j$.

To summarize, the dual adjusted rank of all women $j \in P^t$ such that $w \succeq_m j$, goes down by at least 2. This means that in all subsequent iterations, man m must propose to a woman ranked at or above w in P^t. Eventually, there is a woman, $w' \in P^t$, say, that man m keeps proposing to. This means that the dual adjusted rank of all women in P^t declines by at least 2, while the dual adjusted rank of all women outside of P^t declines by exactly 1. Eventually, some woman outside of P^t must have a dual adjusted rank that is largest and man m will propose to her, a contradiction.

Conclusion

My goal has been to show how one of the most important results in the theory of stable matching can be derived using linear programming. My hope is that it will prompt others to consider similar methods in the analysis of generalizations of the stable matching problem to the many-to-one-case as well as to matching with contracts.

1. Gale, D. and Shapley, L. S., College admissions and the stability of marriage. *Am. Math. Mon.*, 1962, **69**, 9–14.
2. Roth, A. E. and Sotomayor, M., In *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Cambridge University Press, Cambridge, 1990.
3. Roth, A. E., Deferred acceptance algorithms: history, theory, practice and open questions. *Int. J. Game Theory*, 2008, **36**, 537–569.
4. Vande Vate, J. H., Linear programming brings marital bliss. *Oper. Res. Lett.*, 1989, **8**, 147–153.
5. Teo, C. P. and Sethuraman, J., The geometry of fractional stable matching and its applications. *Math. Oper. Res.*, 1998, **23**, 874891
6. Roth, A. E., Rothblum, U. G. and Vande Vate, J. H., Stable matchings, optimal assignments and linear programming. *Math. Oper. Res.*, 1993, **18**, 803–828.

ACKNOWLEDGEMENT. I thank the anonymous referee for many useful suggestions.