Vector spaces: Let $M \in \mathbb{R}^{m \times n}$, and collect the column labels into a set E. Call a subset I of E independent if the corresponding columns form an independent set in $\mathbb{R}^m$.

A couple of observations:

(0) $\emptyset$ is independent

(1) if I is independent, then so is any subset of it

(2) for any $X \subseteq E$, the size of any maximal independent set contained in X is the same

Exercise 1: To prove (2), show the following:

(i) if $I \subseteq X$ is independent, then $|I| \leq \text{rank}(X)$

(ii) if $I \subseteq X$ is independent and $|I| < \text{rank}(X)$, then there exists $c \in X - I$ such that $I \cup \{c\}$ is independent

Graphs: Let $G = (V, E)$. An edge subset $I \subseteq E$ is <u>independent</u> if it contains no cycles.

(0) $\emptyset$ is independent

(1) if I is independent, then so is any subset of it

Equivalently, an edge subset I is independent if it is a forest. As a result, $|I| \leq |V| - \#\text{connected components} = |V| - k(G)$.

Perhaps less trivially, this bound is achieved by any maximal independent set. In fact,

(2) for all $X \subseteq E$, the size of any maximal independent set contained in X is the same (and is equal to $|V(G[X])| - k(G[X])$)

What is a <u>matroid</u>?

Let E be a finite set, called a <u>ground set</u>, and let $\mathcal{I}$ be a family of (distinct) subsets of E, called independent sets. A <u>matroid</u> is a pair $M = (E, \mathcal{I})$, where

(M0) $\emptyset \in \mathcal{I}$

(M1) if $I \in \mathcal{I}$ and $J \subseteq I$ then $J \in \mathcal{I}$

(M2) for any $X \subseteq E$, the size of any $\underbrace{\text{maximal independent set contained in } X}_{= \text{basis for } X}$ is the same $= r(X)$

Some terminology:

- The elements of E are referred to as edges.
- The constant in (M2) is denoted by $r(X)$. In other words, the size of any maximal independent set in X is $r(X)$. The function $r: 2^E \to \{0, 1, 2, \ldots\}$ is called the rank function of M.
- rank of M is $r(M) := r(E)$
- Every independent set of size $r(M)$ is called a basis for M.
- If $X \subseteq E$ is not independent then it is dependent
- Minimal dependent sets are called circuits.

We already saw two examples of a matroid.
The first one is a special case of linear matroids. More generally, to get a linear matroid, all one needs to do is replace $\mathbb{R}$ with any arbitrary field.
The second example was ~~also~~ obtained from a graph G, and it is called the forest matroid of G. If the graph G is not specified, then the matroid is simply referred to as a graphic matroid.

The Greedy Algorithm

Let $M = (E, \mathcal{I})$ be a matroid, cost vector $c \in \mathbb{R}^E$. We are interested in finding a basis of M whose ~~cost~~ (total) cost is maximum:

$$\max\left\{\sum_{e \in B} c_e \ ; \ B \text{ is a basis}\right\}.$$

When M is the forest matroid of a connected graph, then we are looking for a maximum cost spanning tree.

A very natural algorithm for solving this problem is as follows:

The Greedy Algorithm
- Set $I := \emptyset$,
- While there exists an edge $e \in E - I$ such that $I \cup \{e\} \in \mathcal{I}$:
  - pick such e with maximum c_e, and reset $I := I \cup \{e\}$

Theorem 1 [Rado, Edmonds]: Let $M = (E, \mathcal{I})$ be a matroid and $c \in \mathbb{R}^E$ a cost vector. Then the Greedy algorithm finds a maximum cost basis.

Proof: Let B be the output. By the nature of the algorithm, B is a maximal independent set, so it must be a basis, by (M2). Suppose B consists of $e_1, e_2, \dots, e_r$, and they were picked in this order (where $r = \text{rank}(M)$). Note $c_{e_1} \geq c_{e_2} \geq \cdots \geq c_{e_r}$.

Suppose for a contradiction that B is not of maximum cost. So there exists a basis $B' = \{f_1, f_2, \dots, f_r\}$, with $c_{f_1} \geq \cdots \geq c_{f_r}$, whose cost is strictly larger than that of B. There must exist a smallest index k such that $c_{f_k} > c_{e_k}$. So

$$c_{f_1} \geq c_{f_2} \geq \cdots \geq c_{f_k} > c_{e_k}.$$

Since at the k^{th} iteration, we picked e_k and not $f_1, \dots, f_k$, we must have had

$$(*) \left\{ \begin{array}{l} \{e_1, \dots, e_{k-1}, f_1\} \notin \mathcal{I} \quad \text{or} \quad f_1 \in \{e_1, \dots, e_{k-1}\} \\ \quad \vdots \\ \{e_1, \dots, e_{k-1}, f_k\} \notin \mathcal{I} \quad \text{or} \quad f_k \in \{e_1, \dots, e_{k-1}\} \end{array} \right.$$

Let $X = \{e_1, \dots, e_{k-1}, f_1, \dots, f_k\}$. By $(*)$, $\{e_1, \dots, e_{k-1}\}$ is a maximal independent set contained in X. So by (M2), any independent set in X has size at most $k-1$. But $\{f_1, \dots, f_k\} \subseteq B'$ is independent of size k. ∎

Theorem 2 [Borůvka]: Let $(E, \mathcal{I})$ be a pair satisfying (M0) and (M1). Then $(E, \mathcal{I})$ is a matroid if and only if for all $c \in \mathbb{R}^E$, the Greedy algorithm outputs a maximal member of $\mathcal{I}$ of maximum cost.

Proof: Saw $(\Rightarrow)$.

$(\Leftarrow)$: Suppose $(E, \mathcal{I})$ is not a matroid, so it does not satisfy (M2), so there exists $X \subseteq E$ for which there exist two maximal independent sets (members of $\mathcal{I}$) I_1 and I_2 such that $|I_1| > |I_2|$.

Let c be the characteristic vector of X. Then

$$|\text{optimal value}| \geq |I_1|.$$

However, the algorithm could have potentially outputted

$$I_2 \cup \{\text{some edges of } E - X\},$$

whose cost is $|I_2| < |\text{optimal value}|$.

2014 11 1? (TA)

More examples

Uniform Matroids: Take a set E of size n, and choose $k \in \{0,1,\dots,n\}$. Call a subset I of E independent if $|I| \le k$. Let $\mathcal{I}$ be the family of independent sets, and let $U_n^k := (E, \mathcal{I})$.
(M0) and (M1) are obvious. For any $X \subseteq E$, the size of any basis for X is $\min\{|X|, k\} = r(X)$.

Partition Matroid: Let E be a finite set, and let $E_1, \dots, E_r$ be a partition of E. Call a subset $I \subseteq E$ independent if it picks at most one edge from each part, ie $|I \cap E_i| \le 1 \ \forall i \in \{1,\dots,r\}$. Let $\mathcal{I}$ be the family of independent sets. Then $M := (E, \mathcal{I})$ is a matroid.
(M0) and (M1) are obvious. For all $X \subseteq E$, any basis for X has size $|\{i;\ X \cap E_i \neq \emptyset\}| = r(X)$.

Submodularity of the rank function

Let $M = (E, \mathcal{I})$ be a matroid whose rank function is $r: 2^E \to \{0,1,2,\dots\}$.
We will see that r is submodular: ie for all $X, Y \subseteq E$

$$r(X) + r(Y) \ge r(X \cap Y) + r(X \cup Y).$$

Choose $A \subseteq X \cap Y$, $B \subseteq X - Y$, $C \subseteq Y - X$ such that:

- A is a basis for $X \cap Y$;
- $A \cup B$ is a basis for X;
- $A \cup B \cup C$ is a basis for $X \cup Y$.

We have

$$\begin{aligned} r(X \cap Y) + r(X \cup Y) &= |A| + |A \cup B \cup C| \\ &= |A| + |A| + |B| + |C| \\ &= r(X) + |A| + |C| \\ &\le r(X) + r(Y) \end{aligned}$$

because $A \cup C$ is independent and contained in Y.

Exercise: Let E be finite, and let $r: 2^E \to \{0,1,2,\dots\}$ satisfy:
(R0) $r(X) \le |X| \quad \forall X \subseteq E$;
(R1) $r(X) \le r(Y) \quad \forall X \subseteq Y \subseteq E$;
(R2) r is submodular.
Then r is the rank function of a matroid

Minor Operations & the dual matroid

Let $M = (E, \mathcal{I})$ be a matroid whose rank function is r.

Deletion: Take $F \subseteq E$. To delete F is to replace M by

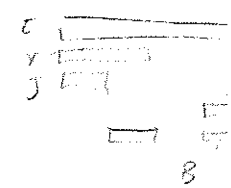
$$M \backslash F := (E-F, \{I \subseteq E-F ; I \in \mathcal{I}\}).$$

For any $X \subseteq E-F$,

$$r_{M \backslash F}(X) = r(X).$$

Duality. Let $\mathcal{I}^*$ be the family of subsets J of E that are *disjoint from* at least one basis of M, ie

$$\mathcal{I}^* := \{J \subseteq E ; r(E-J) = r(E)\}.$$



Claim: $M^* := (E, \mathcal{I}^*)$ is a matroid, and its rank function is

$$r^*(X) = |X| - r(E) + r(E-X)$$

Proof: (M0) and (M1) are ~~also~~ obvious. To see (M2), let $X \subseteq E$. Let J be an M^*-basis for X. Let I be an M-basis for $E-X$. Extend I to a basis B for $M \backslash J$. Since $M \backslash J$ has the same rank as M, B is also a basis for M.

Since J is an M^*-basis for X, we must have $X - J = B \cap X = B - I$.

$$|X| - |J| = |X - J| = |B - I| = |B| - |I| = r(E) - r(E-X)$$

$$\Rightarrow |J| = |X| - r(E) + r(E-X).$$

Thus (M2) holds and M^* is a matroid.

Exercise: Prove that if M is the forest matroid of a (2-connected) planar graph, then M^* is the forest matroid of the plane dual. ? maybe

Contraction: Take $F \subseteq E$. To contract F is to replace M by

$$M/F := (M^* \backslash F)^*.$$

Exercise: Let J be an M-basis for F. Prove that

$$M/F = (E, \{I \subseteq E-F ; I \cup J \in \mathcal{I}\}),$$

and for $X \subseteq E-F$,

$$r_{M/F}(X) = r(X \cup F) - r(F)$$

The Matroid Intersection Theorem: [Edmonds]

Let $M_1=(E,\mathcal{I}_1)$ & $M_2=(E,\mathcal{I}_2)$ be two matroids with rank functions r_1 & r_2. Then

$$\max\{|I|;\ I\in\mathcal{I}_1\cap\mathcal{I}_2\} = \min\{r_1(X)+r_2(E-X);\ X\subseteq E\}.$$

Proof: Take $I\in\mathcal{I}_1\cap\mathcal{I}_2$ and $X\subseteq E$. Then

$$|I| = |I\cap X| + |I-X| \le r_1(X)+r_2(E-X).$$

For the reverse inequality, by induction on $|E|$. Base-case is easy. Let $k :=$.
We are looking for a common independent set of size at least k. Suppose we cannot do so. Then $k\ge 1$. Take an edge $e\in E$ st $\{e\}\in\mathcal{I}_1\cap\mathcal{I}_2$. Such an edge must exist, for if not, let $Y=\{e\in E;\ r_1(e)=0\}$ and write

$$r_1(Y)+r_2(E-Y)=0\ge k\ge 1$$

a contradiction.

Apply induction on $M_1\backslash e$ and $M_2\backslash e$: we get a common independent set of size

$$\ge \min\{r_1(X)+r_2(E'-X);\ X\subseteq E'\}$$

where $E'=E-\{e\}$. By assumption, there must exist a set $A\subseteq E'$ such that

(*) $$r_1(A)+r_2(E'-A)\le k-1.$$

Apply induction on M_1/e and M_2/e. We get a common independent set I of size $\ge$

$$\min\{r_{M_1/e}(X)+r_{M_2/e}(E'-X);\ X\subseteq E'\}$$

Note $I\cup\{e\}$ is also independent for both M_1 and M_2, because $r_1(e)=1$.

So $\exists\, B\subseteq E'$ st

$$r_{M_1/e}(B)+r_{M_2/e}(E'-B)+1\le k-1$$

$$\Vert$$

(**) $$r_1(B\cup\{e\})-1+r_2(E-B)-1+1$$

So (*) and (**) yield

$$2k-1\ge r_1(A)+r_1(B\cup\{e\})+r_2(E'-A)+r_2(E-B)$$
$$\ge r_1(A\cap B)+r_1(A\cup B\cup\{e\})+r_2(E-(A\cap B))+r_2(E-(A\cup B\cup\{e\}))$$

$\ge k \qquad \ge k \qquad \ge 2k.$

Matroid Polytope

$P(M)$ = convex hull of independent sets

$$\chi^I_e = \begin{cases} 1 & \text{if } e \in I, \\ 0 & \text{if } e \in E \setminus I. \end{cases}$$

Theorem: $P(M)$ is defined by

$$\begin{aligned} x(A) &\le r(A) && \forall A \subseteq E, \\ x_e &\ge 0 && \forall e \in E. \end{aligned}$$

Let $w = (w_e : e \in E)$ and consider the max weight indep. set problem. We may assume $w_e \ge 0$ for all $e \in E$. Consider the LP relaxation

$$\begin{aligned} \max \ & \sum (w_e x_e : e \in E) \\ \text{s.t.: } & x(A) \le r(A) \quad \forall A \subseteq E \\ & x_e \ge 0 \quad \forall e \in E \end{aligned}$$

We show solution provided by greedy algorithm is optimal for LP. Order $E = \{e_1, \dots, e_n\}$ such that

$$w_{e_1} \ge w_{e_2} \ge \cdots \ge w_{e_n}.$$

Let x^* be the solution found by the greedy algorithm with elements processed in this order, giving the indep set J. Define

$$T_i = \{e_1, \dots, e_i\}, \quad i \in \{1, \dots, n\}.$$

Dual greedy solution

Dual LP

$$\begin{aligned} \min \ & \sum (r(A) y_A : A \subseteq E) \\ \text{s.t. } & \sum (y_A : e \in A \subseteq E) \ge w_e && \forall e \in E \\ & y_A \ge 0 && \forall A \subseteq E \end{aligned}$$

Set

$$y^*_A = \begin{cases} w_{e_i} - w_{e_{i+1}} & \text{if } A = T_i, \quad 1 \le i \le n-1 \\ w_{e_n} & \text{if } A = T_n \\ 0 & \text{otherwise} \end{cases}$$

Note $y^*_A \ge 0$. Let $e_j \in E$. Then

$$\sum (y^*_A : e_j \in A \subseteq E) = \sum_{i=j}^{n} y^*_{T_i} = \sum_{i=j}^{n-1} (w_{e_i} - w_{e_{i+1}}) + w_{e_n} = w_{e_j}$$

Thus y^* is dual feasible.

CS conditions:

(1) $x_e > 0 \Rightarrow \sum (y_A \,;\, e \in A \subseteq E) = w_e \qquad \forall e \in E$

(2) $y_A > 0 \Rightarrow x(A) = r(A) \qquad \forall A \subseteq E$

(1) holds trivially

(2) $y_A > 0$: Thus A is one of $T_1, \dots, T_n$. We must show $J \cap T_i$ is a basis of T_i. Suppose not. Then $\exists\, e_k \in T_i \setminus J$ such that $(J \cap T_i) \cup \{e_k\} \in \mathcal{I}$. But in iteration k of the greedy algorithm, e_k was not added to J. ✗ ∎

Matroid Intersection

Matroids $M_1 = (E, \mathcal{I}_1)$, $M_2 = (E, \mathcal{I}_2)$.
Want a max-weight common indep set.

Theorem (Edmonds): The convex hull of common indep sets of M_1 and M_2 is

$$(*) \begin{cases} x(A) \le r_1(A) & \forall A \subseteq E \\ x(A) \le r_2(A) & \forall A \subseteq E \\ x_e \ge 0 & \forall e \in E \end{cases}$$

Dual LP

$$\min \sum (r_1(A) y^1_A + r_2(A) y^2_A \,;\, A \subseteq E)$$
$$\text{s.t.: } \sum (y^1_A + y^2_A \,;\, e \in A \subseteq E) \ge w_e \quad \forall e \in E$$
$$y^1_A, y^2_A \ge 0 \qquad \forall A \subseteq E$$

Consider an optimal dual solution $(\bar{y}^1, \bar{y}^2)$. Define w^1, w^2 by

$$w^1_e = \sum (\bar{y}^1_A \,;\, e \in A \subseteq E)$$
$$w^2_e = w_e - w^1_e$$

So $w = w^1 + w^2$.

Dual LP^i

$$\min \sum (r_i(A) y^i_A \,;\, A \subseteq E)$$
$$\text{s.t.: } \sum (y^i_A \,;\, e \in A \subseteq E) \ge w^i_e \quad \forall e \in E$$
$$y^i_A \ge 0 \quad \forall A \subseteq E$$

Lemma: $\bar{y}^i$ is optimal for dual LP^i, $i \in \{1, 2\}$

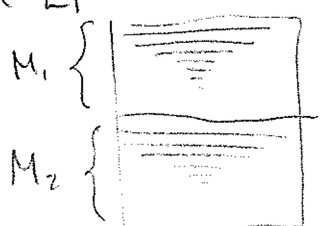
2014 11 17 C

Lemma: If y^{*i} is an optimal solution for dual LP^i, $i \in \{1,2\}$ then (y^{*1}, y^{*2}) is optimal for dual intersection LP.

Theorem: The matroid intersection system (*) is TDI.
Proof: Let $w = (w_e, e \in E)$ be integral. Find opt dual soln and create the split $w = w^1 + w^2$.
Take a greedy dual solution y^{*i} for each dual LP^i, $i \in \{1,2\}$.
So we have two nested families

$$T^1_1 \subset T^1_2 \subset T^1_3 \subset \cdots \subset T^1_{n_1}$$
$$T^2_1 \subset T^2_2 \subset T^2_3 \subset \cdots \subset T^2_{n_2}$$

of subsets such that these are the only positive dual variables.
In LP



- totally unimodular
- reduces to bipartite matching.

2014 11 19

<u>Travelling Salesman Problem (TSP)</u>

$G = (V, E)$ complete graph
$c = (c_e; e \in E)$
$e = uv$, $c_e \equiv$ cost to travel ~~between~~ from u to v.
We study the symmetric TSP, cost is to travel from u to v is the same as from v to u.
TSP: find a min-cost spanning circuit $\equiv$ "Hamiltonian circuit" $\equiv$ "tour"
Reduction from asymmetric to symmetric