MATROIDS
Hereditary Families

Given a **Ground Set** $E$, a **Hereditary Family** $\mathcal{A}$ on $E$ is a collection of subsets $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ (the **independent sets**) such that

\[ I \in \mathcal{I} \text{ and } J \subseteq I \text{ implies that } J \in \mathcal{I}. \]

1. The set $\mathcal{M}$ of matchings of a graph $G = (V, E)$.
2. The set of (edge-sets of) forests of a graph $G = (V, E)$.
3. The set of **stable** sets of a graph $G = (V, E)$. We say that $S$ is stable if it contains no edges.
4. If $G = (A, B, E)$ is a bipartite graph and $\mathcal{I} = \{S \subseteq B : \exists \text{ a matching } M \text{ that covers } S\}$.
5. Let $c_1, c_2, \ldots, c_n$ be the columns of an $m \times n$ matrix $A$. Then $E = [n]$ and $\mathcal{I} = \{S \subseteq [n] : \{c_i, i \in S\} \text{ are linearly independent}\}$. 
An independence system is a matroid if whenever $I, J \in \mathcal{I}$ with $|J| = |I| + 1$ there exists $e \in J \setminus I$ such that $I \cup \{e\} \in \mathcal{I}$. We call this the Independent Augmentation Axiom – IAA.

Matroid independence is a generalisation of linear independence in vector spaces. Only Examples 2, 4 and 5 above are matroids.

To check Example 5, let $A_I$ be the $m \times |I|$ sub-matrix of $A$ consisting of the columns in $I$. If there is no $e \in J \setminus I$ such that $I \cup \{e\} \in \mathcal{I}$ then $A_J = A_I M$ for some $|I| \times |J|$ matrix $M$.

Matrix $M$ has more columns than rows and so there exists $x \neq 0$ such that $Mx = 0$. But then $A_J x = 0$, implying that the columns of $A_J$ are linearly dependent. Contradiction.

These are called Representable Matroids.
To check Example 2 we define the vertex-edge incidence matrix $A_G$ of graph $G = (V, E)$ over $GF_2$.

$A_G$ has a row for each vertex $v \in V$ and a column for each edge $e \in E$. There is a 1 in row $v$, column $e$ iff $v \in e$.

We verify that a set of columns $c_i, i \in I$ are linearly dependent iff the corresponding edges contain a cycle.

If the edges contain a cycle $(v_1, v_2, \ldots, v_k, v_1)$ then the sum of the columns corresponding to the vertices of the cycle is $0$.

To show that a forest $F$ defines a linearly independent set of columns $I_F$, we use induction on the number of edges in the forest. This is trivial if $|E(F)| = 1$. 

MATROIDS
Let $A_F$ denote the submatrix of $A$ made up of the columns corresponding $F$.

Now a forest $F$ must contain a vertex $v$ of degree one. This means that the row corresponding to $v$ in $A_F$ has a single one, in column $e$ say.

Consider the forest $F' = F \setminus \{e\}$. Its corresponding columns $I_{F'}$ are linearly independent, by induction. Adding back $e$ adds a row with a single one and preserves independence. Let $B$ denote $A_{F'}$ minus row $e$.

$$A_F = \begin{bmatrix} 1 & 0 \\ B \end{bmatrix}.$$
We now check Example 4. These are called Transversal Matroids. If $M_1, M_2$ are two matchings in a graph $G$ then $M_1 \oplus M_2 = (M_1 \setminus M_2) \cup (M_2 \setminus M_1)$ consists of alternating paths and cycles.

Suppose now that we have two matchings $M_1, M_2$ in bipartite graph $G = (A, B, E)$. Let $I_j, j = 1, 2$ be the vertices in $B$ covered by $M_j$. Suppose that $|I_1| > |I_2|$.

Then $M_1 \oplus M_2$ must contain an alternating path $P$ with end points $b \in I_1 \setminus I_2, a \in A$. Let $E_1$ be the $M_1$ edges in $P$ and let $E_2$ be the $M_2$ edges of $P$. Then $(M_1 \cup E_1) \setminus E_2$ is a matching that covers $I_1 \cup \{b\}$. 
A matroid is **binary** if it is representable by a matrix over $\mathbb{GF}_2$.

So a graphic matroid is binary.

A matroid is **regular** if it can be represented by a matrix of elements in $\{0, \pm 1\}$ for which every square sub-matrix has determinant $0, \pm 1$. These are called **totally unimodular matrices**.

A matrix with 2 non-zeros in each column, one equal to $+1$ and the other equal to $-1$ is totally unimodular. This implies that graphic matroids are regular. (Take the vertex-edge incidence matrix and replace one of the ones in each column by a -1.)
Partition Matroids

Given a partition $E_1, E_2, \ldots, E_m$ of $E$ and non-negative integers $k_1, k_2, \ldots, k_m$ we define the associated partition matroid as follows:

$$I \in \mathcal{I} \text{ iff } |I \cap E_i| \leq k_i, \ i = 1, 2, \ldots, m.$$ 

Partition matroids are representable.
A matroid **basis** is a maximal independent set i.e. \( B \) is a basis if there does **not** exist an independent set \( I \neq B \) such that \( I \supset B \).

So the bases of the cycle matroid of a graph \( G \) consist of the spanning trees of \( G \).

**Lemma**

*If \( B_1, B_2 \) are bases of a matroid \( \mathcal{M} \), then \(|B_1| = |B_2|\).*

**Proof:** If \(|B_1| > |B_2|\) then there exists \( e \in B_1 \setminus B_2 \) such that \( B_2 \cup \{e\} \) is independent. Contradicting the fact that \( B_2 \) is maximal. □
Theorem

A collection $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ of subsets of $E$ form the bases of a matroid on $E$ iff for all $i, j$ and $e \in B_i \setminus B_j$ there exists $f \in B_j \setminus B_i$ such that $(B_i \cup \{f\}) \setminus \{e\} \in \mathcal{B}$. 

Proof: Suppose first that $\mathcal{B}$ are the bases of a matroid with independent sets $\mathcal{I}$ and that $e \in B_i$ and $e \notin B_j$. Then $B'_i = B_i \setminus \{e\} \in \mathcal{I}$ and $|B'_i| < |B_j|$. So there exists $f \in B_j \setminus B'_i$ such that $B''_i = B'_i \cup \{f\} \in \mathcal{I}$. Now $f \neq e$ since $e \notin B_j$ and $|B''_i| = |B_i|$. So $B''_i$ must be a basis.

Conversely, suppose that $\mathcal{B}$ satisfies the conditions of the theorem and that $\mathcal{I} = \{S : \exists i \text{ s.t. } S \subseteq B_i\}$. Clearly $\mathcal{I}$ is hereditary.
We first argue that all the sets in $\mathcal{B}$ are of the same size.

Suppose that $A = \{i : |B_i| = \max\{|B| : B \in \mathcal{B}\}\}$ and suppose that $A \neq [m]$. Suppose that

$$\min\{|B_i - B_j| : i \in A, j \notin A\} = \left|B_1 \setminus B_2\right|.$$ 

Let $x \in B_1 \setminus B_2$ and let $y \in B_2 \setminus B_1$ be such that $B' = ((B_1 \cup y) \setminus \{x\}) \in \mathcal{B}$.

Then we have $B' \in A$ and $|B' \setminus B_2| < |B_1 \setminus B_2|$, contradiction.
Suppose now that \( I_1, I_2 \in \mathcal{I} \) with \( |I_2| > |I_1| \) and there does not exist \( e \in l_2 \setminus I_1 \) for which \( I_1 \cup \{ e \} \in \mathcal{I} \).

Choose \( B_j \supseteq I_j, j = 1, 2 \) such that \( |B_2 \setminus (I_2 \cup B_1)| \) is minimal.

We must have \( l_2 \setminus B_1 = l_2 \setminus I_1 \). If \( x \in l_2 \cap B_1 \) and \( x \notin I_1 \) then \( I_1 \cup \{ x \} \subseteq B_1 \) and so \( I_1 \cup \{ x \} \in \mathcal{I} \).

Suppose there exists \( x \in B_2 \setminus (I_2 \cup B_1) \). Then by assumption there is \( y \in B_1 \setminus B_2 \) such that \( B' = (B_2 \cup \{ y \}) \setminus \{ x \} \in \mathcal{B} \). But then \( B' \setminus (I_2 \cup B_1) = (B_2 \setminus (I_2 \cup B_1)) \setminus \{ x \} \), contradicting the definition of \( B_2 \).
So $B_2 \subseteq (l_2 \cup B_1) = (l_2 \backslash B_1) \cup (B_1 \backslash l_2) = (l_2 \backslash l_1) \cup (B_1 \backslash l_2)$ and so

$$B_2 \backslash B_1 = l_2 \backslash l_1. \quad (1)$$

We show next that $B_1 \subseteq (l_1 \cup B_2)$. If there exists $x \in B_1 \backslash (l_1 \cup B_2)$ then there exists $y \in B_2 \backslash B_1$ such that $B' = (B_1 \cup \{y\}) \backslash \{x\} \in \mathcal{B}$. But $(l_1 \cup \{x\}) \subseteq B'$, contradiction.

So, $B_1 \backslash B_2 = l_1 \backslash B_2 \subseteq l_1 \backslash l_2$. Since $|B_1 \backslash B_2| = |B_2 \backslash B_1|$ we see from this and (1) that $|l_1 \backslash l_2| \geq |l_2 \backslash l_1|$ and so $|l_1| \geq |l_2|$, contradiction.
If $S \subseteq E$ then its rank

$$r(S) = \max \{|I \in \mathcal{I} : I \subseteq S|\}.$$ 

So $S \in \mathcal{I}$ iff $r(S) = |S|$. We show next that $r$ is submodular.

**Theorem**

If $S, T \subseteq E$ then $r(S \cup T) + r(S \cap T) \leq r(S) + r(T)$.

**Proof:** Let $I_1$ be a maximal independent subset of $S \cap T$ and let $I_2$ be a maximal independent subset of $S \cup T$ that contains $I_2$. (Such a set exists because of the IAA.)

But then

$$r(S \cap T) + r(S \cup T) = |I_1| + |I_2| = |I_2 \cap S| + |I_2 \cap T| \leq r(S) + r(T).$$
For representable matroids this corresponds to the usual definition of rank.

For the cycle matroid of graph $G = (V, E)$, if $S \subseteq E$ is a set of edges and $G_S$ is the graph $(V, S)$ then $r(S) = |V| - \kappa(G_S)$, where $\kappa(G_S)$ is the number of components of $G_S$.

This clearly true for connected graphs and so if $C_1, C_2, \ldots, C_s$ are the components of $G_S$ then $r(S) = \sum_{i=1}^{s} |C_i| - 1 = |V| - s$.

For a partition matroid as defined above,

$$r(S) = \sum_{i=1}^{m} \min\{k_i, |S \cap E_i|\}.$$
A circuit of a matroid $\mathcal{M}$ is a minimal dependent set. If a set $S \subseteq E, S \notin \mathcal{I}$ then $S$ contains a circuit.

So the circuits of the cycle matroid of a graph $G$ are the cycles.

**Theorem**

If $C_1, C_2$ are circuits of $\mathcal{M}$ and $e \in C_1 \cap C_2$ then there is a circuit $C \subseteq (C_1 \cup C_2) \setminus \{e\}$.

**Proof:** We have $r(C_i) = |C_i| - 1$, $i = 1, 2$. Also, $r(C_1 \cap C_2) = |C_1 \cap C_2|$ since $C_1 \cap C_2$ is a proper subgraph of $C_1$.

If $C' = (C_1 \cup C_2) \setminus \{e\}$ contains no circuit then $r(C_1 \cup C_2) \geq r(C') = |C_1 \cup C_2| - 1$. But then

$$|C_1 \cup C_2| - 1 \leq r(C_1 \cup C_2) \leq r(C_1) + r(C_2) - r(C_1 \cap C_2)$$

$$= (|C_1| - 1) + (|C_2| - 1) - |C_1 \cap C_2|.$$

Contradiction.
Theorem

If $B$ is a basis of $\mathcal{M}$ and $e \in E \setminus B$ then $B' = B \cup \{e\}$ contains a unique circuit $C(e, B)$. Furthermore, if $f \in C(e, B)$ then $(B \cup \{e\}) \setminus \{f\}$ is also a basis of $\mathcal{M}$.

Proof: $B' \notin \mathcal{I}$ because $B$ is maximal. So $B'$ must contain at least one circuit.

Suppose it contains distinct circuits $C_1, C_2$. Then $e \in C_1 \cap C_2$ and so $B'$ contains a circuit $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.

But then $C_3 \subseteq B$, contradiction.
**Theorem**

If \( \mathcal{B} \) denotes the set of bases of a matroid \( \mathcal{M} \) on ground set \( E \) then \( \mathcal{B}^* = \{ E \setminus B : B \in \mathcal{B} \} \) is the set of bases of a matroid \( \mathcal{M}^* \), the dual matroid.

**Proof:** Suppose that \( B_1^*, B_2^* \in \mathcal{B}^* \) and \( e \in B_1^* \setminus B_2^* \).

Let \( B_i = E \setminus B_i^*, i = 1, 2 \). Then \( e \in B_2 \setminus B_1 \).

So there exists \( f \in B_1 \setminus B_2 \) such that \( (B_2 \cup \{ e \}) \setminus \{ f \} \in \mathcal{B} \).

This implies that \( (B_2^* \cup \{ f \}) \setminus \{ e \} \in \mathcal{B}^* \). \( \square \)
Suppose that each $e \in E$ is given a weight $w_e$ and that the weight $w(I)$ of an independent set $I$ is given by $w(I) = \sum_{e \in I} c_e$. The problem we discuss is

$$\text{Maximize } w(I) \text{ subject to } I \in \mathcal{I}.$$ 

**Greedy Algorithm:**

begin

Sort $E = \{e_1, e_2, \ldots, e_m\}$ so $w(e_i) \geq w(e_{i+1})$ for $1 \leq i < m$;

$S \leftarrow \emptyset$;

for $i = 1, 2, \ldots, m$;

begin

if $S \cup \{e_i\} \in \mathcal{I}$ then;

begin;

$S \leftarrow S \cup \{e_i\}$;

end;

end;

end
Theorem

The greedy algorithm finds a maximum weight independent set for all choices of $w$ if and only if it is a matroid.

Suppose first that the Greedy Algorithm always finds a maximum weight independent set. Suppose that $\emptyset \neq I, J \in \mathcal{I}$ with $|J| = |I| + 1$. Define

$$w(e) = \begin{cases} 
1 + \frac{1}{2|I|} & e \in I, \\
1 & e \in J \setminus I, \\
0 & e \notin I \cup J.
\end{cases}$$

If there does not exist $e \in J \setminus I$ such that $I \cup \{e\} \in \mathcal{I}$ then the Greedy Algorithm will choose the elements of $I$ and stop. But $I$ does not have maximum weight. Its weight is $|I| + 1/2 < |J|$. So if Greedy succeeds, then the IAA holds.
Conversely, suppose that our independence system is a matroid. We can assume that $w(e) > 0$ for all $e \in E$. Otherwise we can restrict ourselves to the matroid defined by $\mathcal{I}' = \{ I \subseteq E^+ \}$ where $E^+ = \{ e \in E : w(e) > 0 \}$.

Suppose now that Greedy chooses $I_G = e_{i_1}, e_{i_2}, \ldots, e_{i_k}$ where $i_t < i_{t+1}$ for $1 \leq t < k$. Let $I = e_{j_1}, e_{j_2}, \ldots, e_{j_\ell}$ be any other independent set and assume that $j_t < j_{t+1}$ for $1 \leq t < \ell$. We can assume that $\ell \geq k$, for otherwise we can add something from $I_G$ to $I$ to give it larger weight.

We show next that $k = \ell$ and that $i_t \leq j_t$ for $1 \leq t \leq k$. This implies that $w(I_G) \geq w(I)$. 

MATROIDS
Greedy Algorithm

Suppose then that there exists $t$ such that $i_t > j_t$ and let $t$ be as small as possible for this to be true.

Now consider $I = \{e_i : s = 1, 2, \ldots, t - 1\}$ and $J = \{e_j : s = 1, 2, \ldots, t\}$. Now there exists $e_j \in J \setminus I$ such that $I \cup \{e_j\} \in I$.

But $j_s \leq j_t < i_t$ and Greedy should have chosen $e_j$ before choosing $e_{i_t+1}$.

Also, $i_k \leq j_k$ implies that $k = \ell$. Otherwise Greedy can find another element from $I \setminus I_G$ to add.
Minors

Given a graph $G = (V, E)$ and an edge $e$ we can get new graphs by deleting $e$ or contracting $e$.

We describe a corresponding notion for matroids. Suppose that $F \subseteq E$ then we define the matroid $M \backslash F$ with independent sets $\mathcal{I}_{\backslash F}$ obtained by deleting $F$: $I \in I_{\backslash F}$ if $I \in I$, $I \cap F = \emptyset$.

It is clear that the IAA holds for $M \backslash F$ and so it is a matroid.

For contraction we will assume that $F \in I$. Then contracting $F$ defines $M.F$ with independent sets $\mathcal{I}.F = \{I \in I : I \cap F = \emptyset, I \cup F \in I\}$.

We argue next that $M.F$ is also a matroid.
Lemma

$\mathcal{M}.F = (\mathcal{M}_F^*)^*$ and $\mathcal{M}_F = (\mathcal{M}^*.F)^*$.

Proof:

\[
I \in \mathcal{I}.F \iff \exists B \in \mathcal{B}_F, I \subseteq B
\iff \exists B^* \in \mathcal{B}_F^*, I \cap B^* = \emptyset
\iff I \in (\mathcal{I}_F^*)^*.
\]

For the second claim we use

$\mathcal{M}^*.F = (\mathcal{M}_F^{**})^* = (\mathcal{M}_F)^*$.  

□
Suppose we are given two matroids $\mathcal{M}_1, \mathcal{M}_2$ on the same ground set $E$ with $I_1, I_2$ and $r_1, r_2$ etc. having their obvious meaning.

An **intersection** is a set $I \in I_1 \cap I_2$. We give a min-max relation for the size of the largest independent intersection. Let $\mathcal{J}$ denote the set of intersections.

**Theorem (Edmonds)**

$$\max\{J \in \mathcal{J}\} = \min\{r_1(A) + r_2(E \setminus A) : A \subseteq E\}.$$
Before proving the theorem let us see a couple of applications:

Hall’s Theorem: suppose we are given a bipartite graph $G = (A, B, E)$. Let $\mathcal{M}_A, \mathcal{M}_B$ be the following two partition matroids.

For $\mathcal{M}_A$ we define the partition $E_a = \{e \in x : a \in e\}, \ a \in A$. We let $k_a = 1$ for $a \in A$. We define $\mathcal{M}_B$ similarly.

Intersections correspond to matchings and $r_1(A)$ is the number of vertices in $A$ that are incident with an edge of $A$. Similarly $r_2(E \setminus A)$ is the number of vertices in $B$ that are incident with an edge not in $A$. 
For $X \subseteq A$, let

$$A_X = \{v \in A : v \in e \text{ for some } e \in X\}.$$

Define $B_X$ similarly.

So

$$\max\{|M|\} = \min\{|A_X| + |B_{E\setminus X}| : X \subseteq E\}.$$

Now we can assume that if $e \in E \setminus X$ then $e \cap A_X = \emptyset$, otherwise moving $e$ to $X$ does not increase the RHS of the above.

Let $S = A \setminus A_X$. Then $|B_{E\setminus X}| = |N(A)|$ and so

$$\max\{|M|\} = \min\{|A| - |S| + |N(S)| : S \subseteq A\}.$$
Rainbow Spanning Trees: we are given a connected graph $G = (V, E)$ where each edge $e \in E$ is given a color $c(e) \in [m]$ where $m \geq n - 1$. Let $E_i = \{e : c(e) = i\}$ for $i \in [m]$.

A set of edges $S$ is said to be rainbow colored if $e, f \in S$ implies that $c(e) \neq c(f)$.

For a set $A \subseteq E$, we let

$$r_1(A) = c(A) = |\{i \in [m] : \exists e \in A \text{ s.t. } c(e) = i\}|$$

$$r_2(E \setminus A) = n - \kappa(G \setminus A).$$

So, $G$ contains a rainbow spanning tree iff

$$c(A) + (n - \kappa(G \setminus A)) \geq n - 1 \text{ for all } A \subseteq E. \quad (2)$$
We simplify (2) to obtain
\[ c(A) + 1 \geq \kappa(G \setminus A). \] (3)

We can then further simplify (3) as follows: if we add to \( A \) all edges that use a color used by some edge of \( A \) then we do not change \( c(A) \) but we do not decrease \( \kappa(G \setminus A) \).

Thus we can restrict our sets \( A \) to \( E_I = \bigcup_{i \in I} E_i \) for some \( I \subseteq [m] \). Then (3) becomes
\[ \kappa(E_{[m]\setminus I}) \leq |I| + 1 \text{ for all } I \subseteq [m] \]
or
\[ \kappa(E_I) \leq m - |I| + 1 \text{ for all } I \subseteq [m] \]

If you think for a moment, you will see that this is obviously necessary.
Proof of the matroid intersection theorem.

For the upper bound consider $J \in \mathcal{J}$ and $A \subseteq E$. Then

$$|J| = |J \cap A| + |J \setminus A| \leq r_1(A) + r_2(E \setminus A).$$

We assume that $e \in \mathcal{J}$ for all $e \in E$. (Loops can be “ignored”.)

We proceed by induction on $|E|$. Let

$$k = \min\{r_1(A) + r_2(E \setminus A) : A \subseteq E\}.$$

Suppose that $|J| < k$ for all $J \in \mathcal{J}$. 
Then \((M_1)\setminus\{e\}\) and \((M_2)\setminus\{e\}\) have no common independent set of size \(k\). This implies that if \(F = E \setminus \{e\}\) then

\[
r_1(A) + r_2(F \setminus A) \leq k - 1 \quad \text{for some } A \subseteq F.
\]

Similarly, \((M_1).\{e\}\) and \((M_2).\{e\}\) have no common independent set of size \(k - 1\). This implies that

\[
r_1(B) - 1 + r_2(E \setminus (B \setminus \{e\})) - 1 \leq k - 2 \quad \text{for some } e \in B \subseteq E.
\]

This gives

\[
r_1(A) + r_2(E \setminus (A \cup \{e\})) + r_1(B) + r_2(E \setminus (B \setminus \{e\})) \leq 2k - 1.
\]
Matroid Intersection

So, using submodularity and

\[(E \setminus (A \cup \{e\})) \cup (E \setminus (B \setminus \{e\})) = E \setminus (A \cap B)\]

and

\[(E \setminus (A \cup \{e\})) \cap (E \setminus (B \setminus \{e\})) = E \setminus (A \cup B).\]

We have used \(e \notin A\) and \(e \in B\) here. So,

\[r_1(A \cup B) + r_2(E \setminus (A \cup B)) + r_1(A \cap B) + r_2(E \setminus (A \cap B)) \leq 2k - 1.\]

But, by assumption,

\[r_1(A \cup B) + r_2(E \setminus (A \cup B)) \geq k, \quad r_1(A \cap B) + r_2(E \setminus (A \cap B)) \geq k,\]

contradiction.