Trees

A tree is a graph which is

(a) Connected and

(b) has no cycles (acyclic).
Lemma 1 Let the components of $G$ be $C_1, C_2, \ldots, C_r$. Suppose $e = (u, v) \notin E, u \in C_i, v \in C_j$.

(a) $i = j \Rightarrow \omega(G + e) = \omega(G)$.

(b) $i \neq j \Rightarrow \omega(G + e) = \omega(G) - 1$. 

\begin{figure}[h]
\centering
\begin{subfigure}{0.4\textwidth}
\centering
\begin{tikzpicture}
\node[circle, draw] (u) at (0,0) {$u$};
\node[circle, draw] (v) at (2,0) {$v$};
\draw (u) -- (v);
\end{tikzpicture}
\caption{(a)}
\end{subfigure}
\begin{subfigure}{0.4\textwidth}
\centering
\begin{tikzpicture}
\node[circle, draw] (u) at (0,0) {$u$};
\node[circle, draw] (v) at (2,0) {$v$};
\draw (u) -- (v);
\end{tikzpicture}
\caption{(b)}
\end{subfigure}
\end{figure}
Proof Every path $P$ in $G + e$ which is not in $G$ must contain $e$. Also,

$$\omega(G + e) \leq \omega(G).$$

Suppose

$$(x = u_0, u_1, \ldots, u_k = u, u_{k+1} = v, \ldots, u_\ell = y)$$

is a path in $G + e$ that uses $e$. Then clearly $x \in C_i$ and $y \in C_j$.

(a) follows as now no new relations $x \sim y$ are added.

(b) Only possible new relations $x \sim y$ are for $x \in C_i$ and $y \in C_j$. But $u \sim v$ in $G + e$ and so $C_i \cup C_j$ becomes (only) new component.  \[\square\]
Lemma 2 \( G = (V, E) \) is acyclic (forest) with (tree) components \( C_1, C_2, \ldots, C_k \). \(|V| = n\). \( e = (u, v) \not\in E, u \in C_i, v \in C_j \).

(a) \( i = j \Rightarrow G + e \) contains a cycle.

(b) \( i \neq j \Rightarrow G + e \) is acyclic and has one less component.

(c) \( G \) has \( n - k \) edges.
(a) $u, v \in C_i$ implies there exists a path
\((u = u_0, u_1, \ldots, u_\ell = v)\) in $G$.

So $G + e$ contains the cycle $u_0, u_1, \ldots, u_\ell, u_0$. 
Suppose $G + e$ contains the cycle $C$. $e \in C$ else $C$ is a cycle of $G$.

$$C = (u = u_0, u_1, \ldots, u_\ell = v, u_0).$$

But then $G$ contains the path $(u_0, u_1, \ldots, u_\ell)$ from $u$ to $v$ – contradiction.
The drop in the number of components follows from Lemma 1.

The rest of the lemma follows from

(c) Suppose $E = \{e_1, e_2, \ldots, e_r\}$ and $G_i = (V, \{e_1, e_2, \ldots, e_i\})$ for $0 \leq i \leq r$.

Claim: $G_i$ has $n - i$ components.

Induction on $i$.

$i = 0$: $G_0$ has no edges.

$i > 0$: $G_{i-1}$ is acyclic and so is $G_i$. It follows from part (a) that $e_i$ joins vertices in distinct components of $G_{i-1}$. It follows from (b) that $G_i$ has one less component than $G_{i-1}$.

End of proof of claim

Thus $r = n - k$ (we assumed $G$ had $k$ components).
Corollary 1  If a tree $T$ has $n$ vertices then

(a) It has $n - 1$ edges.

(b) It has at least 2 vertices of degree 1, $(n \geq 2)$.

Proof  (a) is part (c) of previous lemma. $k = 1$ since $T$ is connected.

(b) Let $s$ be the number of vertices of degree 1 in $T$. There are no vertices of degree 0 – these would form separate components. Thus

$$2n - 2 = \sum_{v \in V} d_T(v) \geq 2(n - s) + s.$$  

So $s \geq 2$.  \qed
Theorem 1  Suppose \(|V| = n\) and \(|E| = n - 1\). The following three statements become equivalent.

(a) \(G\) is connected.

(b) \(G\) is acyclic.

(c) \(G\) is a tree.

Proof  Let \(E = \{e_1, e_2, \ldots, e_{n-1}\}\) and \(G_i = (V, \{e_1, e_2, \ldots, e_i\})\) for \(0 \leq i \leq n - 1\).
(a) $\Rightarrow$ (b): $G_0$ has $n$ components and $G_{n-1}$ has 1 component. Addition of each edge $e_i$ must reduce the number of components by 1 – Lemma 1(b). Thus $G_{i-1}$ acyclic implies $G_i$ is acyclic – Lemma 2(b). (b) follows as $G_0$ is acyclic.

(b) $\Rightarrow$ (c): We need to show that $G$ is connected. Since $G_{n-1}$ is acyclic, $\omega(G_i) = \omega(G_{i-1}) - 1$ for each $i$ – Lemma 2(b). Thus $\omega(G_{n-1}) = 1$.

(c) $\Rightarrow$ (a): trivial.
Corollary 2  If $v$ is a vertex of degree 1 in a tree $T$ then $T - v$ is also a tree.

Proof  Suppose $T$ has $n$ vertices and $n$ edges. Then $T - v$ has $n - 1$ vertices and $n - 2$ edges. It acyclic and so must be a tree. \hfill \Box
Cut edges

$e$ is a cut edge of $G$ if $\omega(G - e) > \omega(G)$.

**Theorem 2**  
$e = (u, v)$ is a cut edge iff $e$ is not on any cycle of $G$.

**Proof**  
$\omega$ increases iff there exist $x \sim y \in V$ such that all walks from $x$ to $y$ use $e$.

Suppose there is a cycle $(u, P, v, u)$ containing $e$. Then if $W = x, W_1, u, v, W_2, y$ is a walk from $x$ to $y$ using $e$, $x, W_1, P, W_2, y$ is a walk from $x$ to $y$ that doesn’t use $e$. Thus $e$ is not a cut edge.
If $e$ is not a cut edge then $G - e$ contains a path $P$ from $u$ to $v$ ($u \sim v$ in $G$ and relations are maintained after deletion of $e$). So $(v, u, P, v)$ is a cycle containing $e$. □

**Corollary 3**  
A connected graph is a tree iff every edge is a cut edge.
Corollary 4 Every finite connected graph $G$ contains a spanning tree.

Proof Consider the following process: starting with $G$,

1. If there are no cycles – stop.

2. If there is a cycle, delete an edge of a cycle.

Observe that (i) the graph remains connected – we delete edges of cycles. (ii) the process must terminate as the number of edges is assumed finite.

On termination there are no cycles and so we have a connected acyclic spanning subgraph i.e. we have a spanning tree. 

□
Alternative Construction

Let $E = \{e_1, e_2, \ldots, e_m\}$.

begin
  $T := \emptyset$
  for $i = 1, 2, \ldots, m$ do
    begin
      if $T + e_i$ does not contain a cycle
      then $T \leftarrow T + e_i$
    end
  end
  Output $T$
end
Lemma 3  If $G$ is connected then $(V, T)$ is a spanning tree of $G$.

Proof  Clearly $T$ is acyclic. Suppose it is not connected and has components $C_1, C_2, \ldots, C_k$, $k \geq 2$. Let $D = C_2 \cup \cdots \cup C_k$. Then $G$ has no edges joining $C_1$ and $D$ – contradiction. (The first $C_1 : D$ edge found by the algorithm would have been added.)
Theorem 3  Let $T$ be a spanning tree of $G = (V, E)$, $|V| = n$. Suppose $e = (u, v) \in E \setminus T$.

(a) $T + e$ contains a unique cycle $C(T, e)$.

(b) $f \in C(T, e)$ implies that $T + e - f$ is a spanning tree of $G$. 

\begin{center}
\begin{tikzpicture}
  \node (1) at (-7, -3) {f};
  \node (2) at (-7, -1) {v};
  \node (3) at (-4, 0) {u};

  \draw (1) -- (2);
  \draw (2) -- (3);
  \draw (3) -- (1);

  \draw[ dashed] (1) -- (3);

  \begin{scope}[xshift=1cm]
    \node (1) at (-7, -3) {f};
    \node (2) at (-7, -1) {v};
    \node (3) at (-4, 0) {u};

    \draw (1) -- (2);
    \draw (2) -- (3);
    \draw (3) -- (1);
  \end{scope}
\end{tikzpicture}
\end{center}
Proof  (a) Lemma 2(a) implies that $T + e$ has a cycle $C$. Suppose that $T + e$ contains another cycle $C' \neq C$. Let edge $g \in C' \setminus C$. $T' = T + e - g$ is connected, has $n - 1$ edges. But $T'$ contains a cycle $C$, contradicting Theorem 1.

(b) $T + e - f$ is connected and has $n - 1$ edges. Therefore it is a tree. □
Maximum weight trees

$G = (V, E)$ is a connected graph.

$w : E \rightarrow \mathbb{R}$. $w(e)$ is the weight of edge $e$.

For spanning tree $T$, $w(T) = \sum_{e \in T} w(e)$.

Problem: find a spanning tree of maximum weight.
Greedy Algorithm

Sort edges so that $E = \{e_1, e_2, \ldots, e_m\}$ where

$$w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m).$$

begin

$T := \emptyset$

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

begin

if $T + e_i$ does not contain a cycle

then $T \leftarrow T + e_i$

end

end

Output $T$

end

Greedy always adds the maximum weight edge which does not make a cycle with previously chosen edges.
Theorem 4  Let $G$ be a connected weighted graph. The tree constructed by GREEDY is a maximum weight spanning tree.

Proof  Lemma 3 implies that $T$ is a spanning tree of $G$. 

Let the edges of the greedy tree be $e^*_1, e^*_2, \ldots, e^*_{n-1}$, in order of choice. Note that $w(e^*_i) \geq w(e^*_{i+1})$ since neither makes a cycle with $e^*_1, e^*_2, \ldots, e^*_{i-1}$.

Let $f_1, f_2, \ldots, f_{n-1}$ be the edges of any other tree where $w(f_1) \geq w(f_2) \geq \cdots w(f_{n-1})$.

We show that

$$w(e^*_i) \geq w(f_i) \quad 1 \leq i \leq n - 1.$$  (1)
Suppose (1) is false. There exists $k > 0$ such that
\[ w(e_i^*) \geq w(f_i), \ 1 \leq i < k \text{ and } w(e_k^*) < w(f_k). \]

Each $f_i, \ 1 \leq i \leq k$ is either one of or makes a cycle with $e_1^*, e_2^*, \ldots, e_{k-1}^*$. Otherwise one of the $f_i$ would have been chosen in preference to $e_k^*$.

Let components of forest $(V, \{e_1^*, e_2^*, \ldots, e_{k-1}^*\})$ be $C_1, C_2, \ldots, C_{n-k+1}$. Each $f_i, 1 \leq i \leq k$ has both of its endpoints in the same component.
Let \( \mu_i \) be the number of \( f_j \) which have both endpoints in \( C_i \) and let \( \nu_i \) be the number of vertices of \( C_i \). Then

\[
\mu_1 + \mu_2 + \cdots + \mu_{n-k+1} = k \tag{2}
\]
\[
\nu_1 + \nu_2 + \cdots + \nu_{n-k+1} = n \tag{3}
\]

It follows from (2) and (3) that there exists \( t \) such that

\[
\mu_t \geq \nu_t. \tag{4}
\]

[Otherwise

\[
\sum_{i=1}^{n-k+1} \mu_i \leq \sum_{i=1}^{n-k+1} (\nu_i - 1) = \sum_{i=1}^{n-k+1} \nu_i - (n - k + 1) = k - 1.
\]

But (4) implies that the edges \( f_j \) such that \( f_j \subseteq C_t \) contain a cycle. \( \square \)
Cut Sets and Bonds

If $S \subseteq V$, $S \neq \emptyset$, $V$ then the cut-set

$$S : \overline{S} = \{e = vw \in E : v \in S, w \in \overline{S} = V \setminus S\}$$

$$S = \{1, 2, 3\} \quad S : \overline{S} = \{d, e, f\}.$$
Lemma 4  Let $G$ be connected and $X \subseteq E$. Then $G[E \setminus X]$ is not connected iff $X$ contains a cutset.

Proof

Only if

$G[E \setminus X]$ contains components $C_1, C_2, \ldots, C_k$, $k \geq 2$ and so $X \supseteq C_1 : \bar{C}_1$ and $C_1 \neq \emptyset, V$.

If

Suppose $X = S : \bar{S}$ and $v \in S, w \in \bar{S}$. Then every walk from $v$ to $w$ in $G$ contains an edge of $X$.

So $G[E \setminus X]$ contains no walk from $v$ to $w$. \hfill \Box
A **Bond** $B$ is a *minimal* cut-set. I.e. $B = S : \bar{S}$ and if $T : \bar{T} \subseteq B$ then $B = T : \bar{T}$.

$S_1 = \{1, 2, 3\} \quad B_1 = S_1 : \bar{S}_1$ is a bond

$S_2 = \{2, 3, 4, 5\} \quad B_2 = S_2 : \bar{S}_2$ is not a bond

since $B_2 \supset S_3 : \bar{S}_3$ and $B_2 \neq S_3 : \bar{S}_3$ where $S_3 = \{1\}$. 
Theorem 5  \( G \) is connected and \( B \) is a bond \( \iff \) \( G \setminus B \) contains exactly 2 components.

Proof  \( \rightarrow \): \( G \setminus B \) contains components \( C_1, C_2, \ldots, C_k \). Assume w.l.o.g. that there is at least one edge \( e \) in \( G \) joining \( C_1 \) and \( C_2 \). If \( k \geq 3 \) then \( B \supseteq C_3 : \overline{C_3} \) and \( B \neq C_3 : \overline{C_3} \) since \( B \) contains \( e \).

\( \leftarrow \): Assume that \( G \setminus B \) contains exactly two components \( C_1 = G[S], C_2 = G[S] \). Let \( e \in B \). Adding \( e \) to the graph \( C_1 \cup C_2 \) clearly produces a connected graph and so \( B \setminus e \) is not a cutset. \( \square \)
A **co-tree** \( \overline{T} \) of a connected graph \( G \) is the edge complement of a spanning tree of \( G \) i.e. \( \overline{T} = E \setminus T \) for some spanning tree \( T \).

**Theorem 6** Let \( T \) be a spanning tree of \( G \) and \( e \in T \). Then

(a) \( \overline{T} \) contains no bond of \( G \).

(b) \( \overline{T} + e \) contains a unique bond \( B(\overline{T}, e) \) of \( G \).

(c) \( f \in B(\overline{T}, e) \) implies that \( \overline{T} + e - f \) is a co-tree of \( G \).

[Compare with Tree + edge \( \supseteq \) cycle.]
Proof

(a) \( X \subseteq \overline{T} \iff G \setminus X \supseteq T \) which implies that \( G \setminus X \) is connected. So \( X \) is not a bond.

(b)&(c) \( G \setminus (\overline{T} + e) = T \setminus e \) contains exactly two components and so by Theorem 5 \( \overline{T} + e \) contains a bond \( B = S : \overline{S} \) where \( S, \overline{S} \) are the 2 components of \( T \setminus e \).

\[
f \in B \implies e \in C(T, f)
\]
\[
\implies T + f - e \text{ is a tree}
\]
\[
\implies \overline{T} + e - f \text{ is a co-tree proving (c)}
\]

Hence every bond of \( \overline{T} + e \) contains \( f \) - otherwise \( \overline{T} + e - f \) contains a bond, contradicting (a) and proving (b). \( \square \)
$B = (S : \bar{S})$
How many trees? – Cayley’s Formula

n=4

4
12

n=5

5
60
60

n=6

6
120
360
90
360
360

31
Contracting edges

If $e = vw \in E$, $v \neq w$ then we can contract $e$ to get $G \cdot e$ by (i) deleting $e$, (ii) identifying $v, w$ i.e. make them into a single new vertex.

$G - e$ is obtained by deleting $e$.

$\tau(G)$ is the number of spanning trees of $G$. 
Theorem 7  If $e \in E$ is not a loop then

$$\tau(G) = \tau(G \cdot e) + \tau(G - e).$$

Proof

- $\tau(G - e) = \text{the number of spanning trees of } G \text{ which do not contain } e.$

- $\tau(G \cdot e) = \text{the number of spanning trees of } G \text{ which contain } e.$

[Bijection $T \to T \cdot e$ maps spanning trees of $G$ which contain $e$ to spanning trees of $G \cdot e.$]
Matrix Tree Theorem

Define the $V \times V$ matrix $L = D - A$ where $A$ is the adjacency matrix of $G$ and $D$ is the diagonal matrix with $D(v, v) = \text{degree of } v$.

\[
L = \begin{pmatrix}
3 & -1 & -1 & -1 \\
-1 & 3 & -1 & -1 \\
-1 & -1 & 3 & -1 \\
-1 & -1 & -1 & 3 \\
\end{pmatrix}
\]

\[
L_1 = \begin{pmatrix}
3 & -1 & -1 \\
-1 & 3 & -1 \\
-1 & -1 & 3 \\
\end{pmatrix}
\]

Determinant $L_1 = 16$

Let $L_1$ be obtained by deleting the first row and column of $L$.

Theorem 8

\[\tau(G) = \text{determinant } L_1.\]
There is a 1-1 correspondence $\phi_V$ between spanning trees of $K_V$ (the complete graph with vertex set $V$) and sequences $V^{n-2}$. Thus for $n \geq 2$

$$\tau(K_n) = n^{n-2}$$

Cayley’s Formula.

Assume some arbitrary ordering $V = \{v_1 < v_2 < \cdots < v_n\}$.

$\phi_V(T)$:

begin

$T_1 := T$;

for $i = 1$ to $n - 2$ do

begin

$s_i :=$ neighbour of least leaf $\ell_i$ of $T_i$.

$T_{i+1} = T_i - \ell_i$.

end

$\phi_V(T) = s_1 s_2 \cdots s_{n-2}$

end
6,4,5,14,2,6,11,14,8,5,11,4,2
Lemma 5 \[ v \in V(T) \text{ appears exactly } d_T(v) - 1 \text{ times in } \phi_V(T). \]

Proof \quad \text{Assume } n = |V(T)| \geq 2. \text{ By induction on } n.

\( n = 2: \phi_V(T) = \Lambda = \text{empty string}. \)

\( n \geq 3: \)

\[ \phi_V(T) = s_1 \phi_{V_1}(T_1) \text{ where } V_1 = V - \{\ell_1\}. \]

\( s_1 \text{ appears } d_{T_1}(s_1) - 1 + 1 = d_T(s_1) - 1 \text{ times} \text{ - induction.} \)

\( v \neq s_1 \text{ appears } d_{T_1}(v) - 1 = d_T(v) - 1 \text{ times} \text{ - induction.} \)

\( \square \)

37
Construction of $\phi_V^{-1}$

Inductively assume that for all $|X| < n$ there is an inverse function $\phi_X^{-1}$. (True for $n = 2$).

Now define $\phi_V^{-1}$ by

$$\phi_V^{-1}(s_1s_2\ldots s_{n-2}) = \phi_{V_1}^{-1}(s_2\ldots s_{n-2}) \text{ plus edge } s_1\ell_1,$$

where $\ell_1 = \min\{s : s \notin s_1, s_2, \ldots s_{n-2}\}$ and $V_1 = V - \{\ell_1\}$.

Then

$$\phi_V(\phi_V^{-1}(s_1s_2\ldots s_{n-2})) =$$

$$= \phi_V(\phi_{V_1}^{-1}(s_2\ldots s_{n-2}) \text{ plus edge } s_1\ell_1)$$

$$= s_1\phi_{V_1}(\phi_{V_1}^{-1}(s_2\ldots s_{n-2}))$$

$$= s_1s_2\ldots s_{n-2}.$$

Thus $\phi_V$ has an inverse and the correspondence is established.
Number of trees with a given degree sequence

Corollary 5  If \(d_1 + d_2 + \cdots + d_n = 2n - 2\) then the number of spanning trees of \(K_n\) with degree sequence \(d_1, d_2, \ldots, d_n\) is

\[
\binom{n-2}{d_1-1,d_2-1,\ldots,d_n-1} = \frac{(n - 2)!}{(d_1 - 1)!(d_2 - 1)!\cdots(d_n - 1)!}.
\]

Proof  From Pfuffer’s correspondence and Lemma 5 this is the number of sequences of length \(n - 2\) in which 1 appears \(d_1 - 1\) times, 2 appears \(d_2 - 1\) times and so on. \(\square\)