Notes for Math 446, Fall 2004 Lehigh University: some basic polyhedral results, total unimodularity, Konig-Egevary Theorem.

These notes cover material covered in class not in the text and material in the text that we proved differently in class (usually because the text assumed previous results that we had not covered). For even more details see the suggested extra references.

A few basics about polyhedra

A polyhedron in $\mathbf{R}^n$ is defined to be $P = \{\boldsymbol{x} | A\boldsymbol{x} \leq \boldsymbol{b}\}$ for some $m \times n$ matrix A and vector $\boldsymbol{b}$ in $\mathbf{R}^m$. That is, it is the intersection of a finite number of half spaces. A basic result in the theory of polyhedra is the Weyl-Minkowski Theorem (which we will not prove here) which states that P is a polyhedron in $\mathbf{R}^n$ if and only if there exist finite sets V, R (of vectors in $\mathbf{R}^n$) such that $\boldsymbol{x} \in P$ if and only if there exist scalers $\lambda_{\boldsymbol{v}}, \mu_{\boldsymbol{r}}$ such that

$$\boldsymbol{x} = \sum_{\boldsymbol{v} \in V} \lambda_{\boldsymbol{v}} \boldsymbol{v} + \sum_{\boldsymbol{r} \in R} \mu_{\boldsymbol{r}} \boldsymbol{r} \text{ with } \sum_{\boldsymbol{v} \in V} \lambda_{\boldsymbol{v}} = 1 \text{ and } \lambda_{\boldsymbol{v}}, \mu_{\boldsymbol{r}} \geq 0.$$

Informally, the polyhedron is described by a set of extreme points (vertices) V and extreme rays R. A bounded polyhedron is called a polytope and the bounded version of the theorem above states that a polytope is the convex hull of a finite set V of points. (The convex hull of a set V is the set of all points that are convex combinations of the points in V where a a convex combination satisfies the conditions for the λ above.)

A *supporting hyperplane* for a polyhedron P is a hyperplane $\{\boldsymbol{x} | \boldsymbol{c}\boldsymbol{x} = M\}$ for which $\max\{\boldsymbol{c}\boldsymbol{x} | x \in P\}$ is finite. A *face* is the intersection of P with a supporting hyperplane. So a face is the set of optimal solutions to a linear programming problem over P for some $\boldsymbol{c}$. A point that is a face is called a vertex or an extremal point.

Proposition: The following are equivalent for a polyhedron $P = \{\boldsymbol{x} | A\boldsymbol{x} \leq \boldsymbol{b}\}$:
(i) F is a face of P
(ii) There exists a vector $\boldsymbol{c}$ such that $M = \max\{\boldsymbol{c}\boldsymbol{x} | x \in P\}$ is finite and $F = \{\boldsymbol{x} \in P | \boldsymbol{c}\boldsymbol{x} = M\}$.
(iii) $F = \{\boldsymbol{x} \in P | A'\boldsymbol{x} = \boldsymbol{b}'\} \neq \emptyset$ for some subsystem $A'\boldsymbol{x} \leq \boldsymbol{b}'$ of $A\boldsymbol{x} \leq \boldsymbol{b}$.

Proof outline: The equivalence of (i) and (ii) us essentially the definition of face as described above.

(ii) $\Rightarrow$ (iii): Assuming (ii), by duality $M = \min\{\boldsymbol{y}\boldsymbol{b} | \boldsymbol{y}A = \boldsymbol{c}, \boldsymbol{y} \geq \boldsymbol{0}\}$. Let $\boldsymbol{y}^*$ be a dual optimal solution and let $A'\boldsymbol{x} \leq \boldsymbol{b}'$ be the subsystem of $A\boldsymbol{x} \leq \boldsymbol{b}$ consisting of the inequalities with corresponding entry in $\boldsymbol{y}^*$ positive. Then $F = \{\boldsymbol{x} \in P | A'\boldsymbol{x} = \boldsymbol{b}'\}$, since if $A\boldsymbol{x} \leq \boldsymbol{b}$ then $\boldsymbol{c}\boldsymbol{x} = M \Leftrightarrow \boldsymbol{y}^* A\boldsymbol{x} = \boldsymbol{y}^*\boldsymbol{b} \Leftrightarrow A'\boldsymbol{x} = \boldsymbol{b}'$

(iii) $\Rightarrow$ (ii): If $F = \{\boldsymbol{x} \in P | A'\boldsymbol{x} = \boldsymbol{b}'\} \neq \emptyset$, let $\boldsymbol{c}$ be the sum of the rows of A' and let M be the sum of the components of $\boldsymbol{b}'$. Then $\boldsymbol{c}\boldsymbol{x} \leq M$ for all $\boldsymbol{x} \in P$ and $F = \{\boldsymbol{x} \in P | \boldsymbol{c}\boldsymbol{x} = M\}$. $\square$

From this we can conclude that P has only finitely many faces. In addition, we see that a vertex arises as a solution to a system of equalities. This will allow us to use Cramer's rule in our results about total unimodularity below.

Total unimodularity

Recall that a totally unimodular matrix (TUM) is a matrix with entries from $\{0, +1, -1\}$ such that the determinant of every square submatrix is $0, +1$ or -1.

As in the text it is easy to see the following using basic properties of determinants:

Lemma: If A is TUM then so are $[A|I]$, A^T and $[A| - A]$.

We will use this lemma in the proof below. It also allows the extensions of the Hoffman-Kruskal Theorem to systems with integral upper and lower bounds on the vertices.

Hoffman-Kruskal Theorem: Let A be an integral matrix. Then A is totally unimodular if and only if the polyhedron $\{\boldsymbol{x}|A\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$ is integral.

We will omit the proof of sufficiency.

Proof outline for necessity: The minimal faces of $P = \left[\frac{A}{-I}\right] \boldsymbol{x} \leq \left[\frac{\boldsymbol{b}}{\boldsymbol{0}}\right]$ are vertices. This follows from the fact that $\left[\frac{A}{-I}\right]$ has rank n if A is $m \times n$ and the fact (which we do not prove) that minimal faces have dimension $n - Rank(A)$.)

Suppose that A is totally unimodular. Let $\boldsymbol{b}$ be some integral vector and $\boldsymbol{x}^*$ a vertex of $\boldsymbol{P}$. From the result above, $\boldsymbol{x}^*$ is the solution of $\boldsymbol{A'x} = \boldsymbol{b'}$ for some subsystem of $\left[\frac{\boldsymbol{A}}{-\boldsymbol{I}}\right] \boldsymbol{x} \leq \left[\frac{\boldsymbol{b}}{\boldsymbol{0}}\right]$ with $\boldsymbol{A'}$ a nonsingular $\boldsymbol{n} \times \boldsymbol{n}$ matrix. Since $\boldsymbol{A}$ is totally unimodular, $\boldsymbol{A'}$ has determinant $+\boldsymbol{1}$ or $-\boldsymbol{1}$. Then by Cramer's rule $\boldsymbol{x}^* = (\boldsymbol{A'})^{-1}\boldsymbol{b}$ is integral. $\square$

Konig's Theorem and Konig-Egevary Theorem

Recall that a bipartite graph is a graph whose vertex set can be partitioned into two parts $\boldsymbol{A}, \boldsymbol{B}$ such that every edge has one end in each part. Recall also the definitions in section 6.1 of the text. The notation is

maximum size of an independent set	$\boldsymbol{\alpha(G)}$
maximum size of a matching	$\boldsymbol{\alpha'(G)}$
minimum size of a vertex cover of edges	$\boldsymbol{\beta(G)}$
minimum size of an edge cover of vertices	$\boldsymbol{\beta'(G)}$

We will give here a direct linear programming proof of the Konig-Egevary Theorem and Konig's Theorem as well as the direct proof we covered in class. For other proofs see section 6.1 of the text. This can also be proved from network or circulation results.

Observe that for general graphs we have max independent set of edges (matching) $\leq$ min vertex cover of the edges of a maximum matching are independent so no vertex can cover more than one edge of a matching. Similarly, for graphs with no isolated vertices (so that edge covers are defined) we have max independent set of vertices $\leq$ min edge cover of vertices . We will assume this in what follows.

Konig-Egevary Theorem: In a bipartite graph max independent set of edges (matching) $=$ min vertex cover of edges.

Konig's Theorem: In a bipartite graph with no isolated vertices max independent set of vertices $\leq$ min edge cover of vertices .

We first give a short direct proof of the Konig-Egevary Theorem by Rizzi (2000). There are many other direct proofs.

Proof 1 of Konig-Egevary: Let G be a minimal counterexample. Then G is connected and is not a cycle or a path. (Note, here is where we use the bipartite assumption as the statement is not true for odd cycles.) A graph with every vertex degree at most 2 has every component a path or a cycle. (This is not hard to show. We will assume it.) Thus G has a vertex u with degree at least 3. Let v be some vertex adjacent to u. If $\alpha'(G-v) < \alpha'(G)$ then, by minimality of G, $G-v$ has a cover W' with $|W'| = \alpha'(G-v) < \alpha'(G)$. Then $W' \cup \{v\}$ is a vertex cover of G with size $\alpha'(G)$.

Thus, we may assume that there exists a maximum matching M of G with no edge incident to v. Let f be an edge of $G-M$ incident at u but not at v which exists as u has degree at least 3. By minimality $G-f$ has a cover W' with $|W'| = \alpha'(G-f) = \alpha(G)$. Since no edge of M is incident to v, it follows that W' does not contain v. (Since each vertex of W' covers at most one edge of M and $|W'| = |M|$ every vertex must cover some edge of M. So v cannot be in W'.) With $v \notin W'$ and edge uv covered we must have $u \in W'$. Since u covers edge f, W' is a cover of G with size $\alpha'(G)$. This contradicts the assumption of a minimal counterexample. $\square$

We now show how both the Konig-Egevary Theorem as well as Konig's Theorem, weighted generalizations of them as well as other generalizations follow directly from linear programming duality and total unimodularity.

Total unimodularity of bipartite incidence matrices: If A is the vertex-edge incidence matrix of a bipartite graph the A is totally unimodular. This follows from Ghouila-Houri's condition for total unimodularity. Partition any set of rows into those corresponding to vertices in bipartite part A and those in bipartite part B. Then each column has either one or no 1's in each part and total unimodularity follows from the condition.

Let A be the incidence matrix of a graph. Consider the linear programming problems (independent edges) $\max\{1x | Ax \leq 1, x \geq 0\}$ and (vertex covering) $\min\{y1 | yA \geq 1, y \geq 0\}$. As discussed in the beginning of section 19.3 (there in the more general setting of hypergraphs), the optimal *integral* solutions to these problems correspond to maximum independent sets of edges (maximum matching) and minimum vertex cover of edges respectively. (The constraints for the matching problem force an integral solution to have all entries 0 or 1. It is also not hard to see that an optimal integral solution to the covering problem will also have entries 0 or 1 as there is no need for greater integers to satisfy the constraints.) For general graphs (and hypergraphs) the optimal solutions are not necessarily integral. Duality in these cases assert that the maximimum fractional matching size equals the minimum fractional covering, where the fractional values can be interpreted as the solutions to the LPs.

When the graph is bipartite, total unimodularity of the incidence matrix ensures that there are integral optimal solutions. This along with duality proves the Konig-Egevary Theorem.

Now consider the linear programming problems (independent vertices) $\mathbf{max\{1}\boldsymbol{x}|\boldsymbol{A^T x} \leq \mathbf{1}, \boldsymbol{x} \geq \mathbf{0\}}$ and (edge covering) $\mathbf{min\{}\boldsymbol{y}\mathbf{1}|\boldsymbol{yA^T} \geq \mathbf{1}, \boldsymbol{y} \geq \mathbf{0\}}$. These are the packing and covering problems discussed in the beginning of section 19.1 (there in the more general setting of hypergraphs), the optimal *integral* solutions to these problems correspond to maximum independent sets of vertices and minimum edge covers of vertices respectively. (In the hypergraph setting interchanging the role of vertices and edges corresponds to taking the transpose of the incidence matrix so in some sense the independent edges/vertex covering problems and the independent vertices/edge covering are the same, just with the terms edges and vertices interchanged.)

Here when $\boldsymbol{A}$ is the incidence matrix of a bipartite graph, total unimodularity and duality ensure integral solutions. This is Komig's Theorem.

In each of these cases we can generalize to replace the two different $\mathbf{1}$ vectors with $\boldsymbol{c}$ and $\boldsymbol{b}$ to obtain weighted versions of the Konig-Egevary Theorem and Konig's Theorem.

There are a number of closely related optimization problems for which results follow from total unimodularity of the bipartite incidence matrix. For example, the *assignment problem* has $\boldsymbol{n}$ workers and $\boldsymbol{n}$ jobs and a utility for each worker-job pairing. The assignment problem is to find a maximum utility pairing of the workers and jobs. This is $\mathbf{max\{}\boldsymbol{cx}|\boldsymbol{Ax} = \mathbf{1}, \boldsymbol{x} \geq \mathbf{0\}}$ where $\boldsymbol{A}$ is the incidence matrix of a complete (all possible bipartite edges present) bipartite graph with both parts having the same size. The Hitchcock-Koopmans transportation problem is a special case of the transportation problem in section 19.2 where the network is a bipartite graph with supplies in one part and demands in the other. Historically, Hitchcock, Koopmans and Kantorovich all examined variations on these sorts of transportation problems around 1940 and were precursors to the development of linear programming by Dantzig. Koopmans and Kantorovich later one the Nobel prize in economics for their work.

Dilworth's Theorem

One final comment in these notes is a reminder of our discussion of the network based proof in the text of Dilworth's Theorem. Recall that the theorem invoked is not the max-flow min-cut theorem but the min-flow max-demand theorem which is mentioned in the text but not discussed extensively. It can be proved as the max-flow min-cut theorem, using the algorithm developed in the text, or in a manner similar to the homework proof of max-flow min-cut from the circulation theorem, or directly from duality and total unimodularity of the incidence matrix of a digraph as discussed in class and done in the text.