Essential Coding Theory — Madhu Sudan
6.896
Due: Wednesday, September 11, 2002

Problem Set 1

Instructions

References: In general, try not to run to reference material to answer questions. Try to think about the problem to see if you can solve it without consulting any external sources. If this fails, you may look up any reference material.

Collaboration: Collaboration is allowed, but limit yourselves to groups of size at most four.

Writeup: You must write the solutions in latex, by yourselves. Cite all references and collaborators. Explain why you needed to consult any of the references, if you did consult any.

Problems

1. (Linear Algebra Review): **(Need not be turned in.)**

   (a) Given a $k \times n$ matrix G with 0/1 entries, of rank k over $\mathbb{Z}_2$, generating a linear code $C = \{\mathbf{x} \cdot G | \mathbf{x}\}$, show that there exists an $n \times m$ matrix H, (henceforth referred to as the parity check matrix), such that $C = \{\mathbf{y} | \mathbf{y}H = \mathbf{0}\}$. What is the relationship between m, n and k above?

   By elementary linear algebra involving row operations (replacing row i by row i plus row j), and column exchanges, we can write G as $(I_k|A)$ where A is some $k \times n-k$ matrix. Now the matrix $H = \begin{pmatrix} -A \\ I_{n-k} \end{pmatrix}$ satisfies $GH = 0$. Thus every vector $\mathbf{y} = \mathbf{x}G$ satisfies $\mathbf{y}H = \mathbf{0}$. Furthermore standard linear algebra implies that the space of vectors that satisfies $\mathbf{y}H = 0$ is at most k and so this must be the space spanned by G. The relationship between n, k, m is thus $m = n - k$.

   (b) Give an efficient algorithm to compute such an H, given G, and vice versa.

   Essentially Guassian Elimination. Writing the code is tedious.

   (c) Give an explicit description of the generator matrix of a Hamming code of block length $2^\ell - 1$.

   Lets permute the coordinates of the Hamming code, so that the parity check matrix has rows of non-increasing weight (and so the last ℓ rows form the identity matrix). The generator corresponding to this permutation of the coordinates has as its rows vectors of the form $\langle \mathbf{e}_i, \mathbf{b}_i \rangle$ where for $i \in [2^\ell - \ell - 1]$, the vectors $\mathbf{e}_i$'s are vectors of length $2^\ell - \ell - 1$ with the ith vectors being 1 exactly in the ith coordinate, and $\mathbf{b}_i$'s are all vectors of length ℓ with at least two coordinates being 1's.

2. (Binary Hamming code & bound):

(a) What is the rate of the Hamming code of block length $2^\ell - 1$?

By Problem 1.(a), the message length of the code is $2^\ell - \ell - 1$ and so the rate is $1 - \frac{\ell}{2^\ell - 1}$.

(b) Show that if C is a t-error-correcting code in $\{0,1\}^n$, then $|C| \leq 2^n / \mathrm{Vol}(n,t)$, where $\mathrm{Vol}(n,t) = \sum_{i=0}^{t} \binom{n}{i}$.

By definition, the Hamming balls of radius t around codewords are non-intersecting in a t-error correcting code. Since each such Hamming ball is a subset of $\{0,1\}^n$ we get that the sum of their volumes is at most 2^n. Each has volume equal to $\mathrm{Vol}(n,t)$ and this gives the bound.

(c) Conclude that the Hamming codes of Part (a) are optimal in their performance.

Hamming codes are 1-error correcting codes. The bound of Part (b) implies such codes may have at most $2^n/(n+1)$ codewords. Part (a) shows that Hamming codes do achieve this bound exactly when $n = 2^{\ell-1}$. (The number of codewords is $2^{2^\ell - \ell - 1} = 2^{n-\ell} = 2^n/2^\ell = 2^n/(n+1)$.)

3. (Extra Credit Question) For general q, give the best construction you can of a q-ary code of minimum distance 3.

When q is a prime power, we can let Σ the alphabet be $\mathbb{F}_q$ a finite field of size q. We can then pick H, the parity check matrix, to be all non-zero vectors in $\mathbb{F}_q^\ell$ with the first non-zero entry being 1 (so no vector is zero and no two are scalar multiples of each other). This gives an $n \times \ell$ parity check matrix with $n = (q^\ell - 1)/(q-1)$ and thus a $[n = (q^\ell - 1)/(q-1), n - \ell, 3]_q$ code. This can be show to be optimal as in Problem 2.
I don't know what happens if q is not a prime power. I'd be interested to find out!

4. (Pairwise independent spaces):

(a) Let H be the $(2^\ell - 1) \times \ell$ parity check matrix of a binary Hamming code. Show that the collection of column vectors $\{H\mathbf{x}^T | \mathbf{x} \in \{0,1\}^\ell\}$ forms a pairwise independent space.

This is a special case of a more general result. If C is a code of minimum distance d, then any $d-1$ rows of its parity check matrix H are linearly independent. This implies that the projection of $H\mathbf{x}^T$ to any $d-1$ coordinates is random when $\mathbf{x}$ is random. (This is worked out in greater detail below, where M is supposed to be the $d-1$ rows of H that we are focussing on.) Thus the set of vectors $\{H\mathbf{x}^T | \mathbf{x}\}$ forms a $(d-1)$-wise independent space. Using the fact that Hamming codes have $d = 3$, gives us the pairwise independent case.
Claim: If M is a $(d-1) \times \ell$ matrix over $\mathbb{F}_q$ of rank $d-1$, then $M\mathbf{x}^T$ is random if $\mathbf{x}$ is chosen randomly from $\mathbb{F}_q^\ell$.
Proof: Write M as $(A|B)$ where A is a square matrix of full rank. Write $\mathbf{x}$ correspondingly as $\langle \mathbf{x}_1 \mathbf{x}_2 \rangle$. Now consider the probability that $M\mathbf{x}^T = \mathbf{a}$ for a fixed $\mathbf{a} \in \mathbb{F}_q^{d-1}$. This is equivalent to the condition that $\mathbf{x}_1 = A^{-1}(\mathbf{a} - B\mathbf{x}_2)$ which happens with probability exactly $q^{-(d-1)}$.

(b) (Extra Credit Question) Show that any pairwise independent space on n bits must contain at least $n+1$ points.

Sketch: (We'll see rigorous proofs that follow the outline below, later in the course. Right now, you should at least scan the proof to get the gist of it.)
Let's write bits as $+1$ or -1. So the n bit vectors in the sample space now become vectors in $\{+1,-1\}^n$. Say we have m such vectors $\mathbf{v}_1,\ldots,\mathbf{v}_m$. Consider the $m \times n$ matrix M with the vectors $\mathbf{v}_i$'s as its rows. Let $\mathbf{c}_1,\ldots,\mathbf{c}_n$ be its columns. Augment the $\mathbf{c}_j$'s with the vector $\mathbf{c}_0$ which is a 1 in every entry. From the one-wise independence of the vectors we have that $\mathbf{c}_0$ is orthogonal to the remaining $\mathbf{c}_j$'s, when viewed as vectors over the reals. From the pairwise independence we now have that $\mathbf{c}_j$'s are pairwise orthogonal too. Thus we have $n+1$ vectors in m-dimensional real space, which are pairwise orthogonal. Standard linear algebra implies these are the number of dimensions is at least the number of vectors, i.e., $m \geq n+1$.

5. A Directed Cut (DiCut) in a directed graph $G=(V,E)$ is an ordered partition $(S,\overline{S})$ of V. The size of the DiCut is the number of edges $(u,v) \in E$ with $u \in S$ and $v \in \overline{S}$.

   (a) Show that every graph has a DiCut of size at least $|E|/4$.

   Pick a random partition $(S,\overline{S})$ of G, i.e., each vertex $u \in V$ decides independently with probability half whether it wants to be in S or $\overline{S}$. The probability that a given edge is in the DiCut is $\frac{1}{4}$. Thus, by linearity of expectations, the expected number of edges in this random cut is $\frac{|E|}{4}$. In particular there exists a cut with $\frac{|E|}{4}$ edges.

   (b) Give a deterministic polynomial time algorithm to find such a DiCut in a given graph.

   (There are two natural solutions to this problem - one that involves pairwise independence and one that doesn't. Guess which one I want.)

   Let X_u denote the bit corresponding to choice of vertex u. The analysis in Part (a) continues to work if the choices X_u are pairwise independent. Since we know (by Problem 4.(a) and the properties of the Hamming code) that pairwise independent spaces of vectors exist in $\{0,1\}^n$ with $n+1$ sample points, we can pick such a space and we know for one vector of choices $\langle X_u \rangle_u$ in this space, the choices give a cut with value at least the expectation, i.e. with at least $|E|/4$ edges in the DiCut.

6. The Hat Problem:

   (a) Lets say that a directed graph G is a subgraph of the n-dimensional hypercube if its vertex set is $\{0,1\}^n$ and if $u \to v$ is an edge in G, then u and v differ in at most one coordinate. Let $K(G)$ be the number of vertices of G with in-degree at least one, and out-degree zero. Show that the probability of winning the hat problem equals the maximum, over directed subgraphs G of the n-dimensional hypercube, of $K(G)/2^n$.

   Assume w.l.o.g. that we only consider graphs which do not contain both the edge pairs $u \to v$ and $v \to u$, since this graph does not have a larger $K(G)$ than the graph in which both these edges are deleted.
   We can now draw a 1-1 correspondence between strategies for guessing and subgraphs of the hypercube (provided the subgraph does not contain edges of the form $u \to v$ and $v \to u$). The vertices correspond to the assignment of the hats, and the unordered pair of vertices $\{v,u\}$ where v and u differ in only the ith coordinate corresponds to the view of the ith player. If on this view the player guesses that u is the right view, then lets draw an edge $v \to u$, if the player guess that v is the

right view, lets draw an edge from $u \to v$ and lets not draw any edges between u and v if the player abstains.
In the graph obtained this way every vertex with positive indegree and zero outdegree corresponds to a winning position. Thus the winning probability corresponding to this strategy is $K(G)/2^n$.

(b) Using the fact that the out-degree of any vertex is at most n, show that $K(G)/2^n$ is at most $\frac{n}{n+1}$ for any directed subgraph G of the n-dimensional hypercube.

Let S be the set of vertices with positive in-degree and out-degree zero in G. I.e. $|S| = K(G)$. We have $\sum_{v\in\{0,1\}^n} \text{in-deg}(v) \geq \sum_{v\in S} \text{in-deg}(v) \geq |S| = K(G)$. We have $\sum_{v\in\{0,1\}^n} \text{out-deg}(v) = \sum_{v\notin S} \text{out-deg}(v) \leq n(2^n - |S|)$. And finally $\sum_v \text{in-deg}(v) = \sum_v \text{out-deg}(v)$. Thus we get $K(G) \leq n(2^n - K(G)) \Leftrightarrow K(G)/2^n \leq \frac{n}{n+1}$ for any graph G.

(c) Show that if $n = 2^\ell - 1$, then there exists a directed subgraph G of the n-dimensional hypercube with $K(G)/2^n = \frac{n}{n+1}$. (This is where the Hamming code comes in.)

Let $C \subseteq \{0,1\}^n$ be any code of distance at least 3. Construct G as follows: For every pair of vertices such that $u \in C$ and $v \notin C$ such that $\Delta(u,v) = 1$, draw an edge $u \to v$. Since the distance of the code is at least 3, this ensures there are no edges from $u \to v$ and $v \to u$. Furthermore, every vertex at distance 1 from a codeword has out-degree zero and so $K(G) = n \cdot |C|$. To optimize this construction, we need a code of distance 3 which has maximum number of codewords. Hamming codes give this to us, when they exist, with $|C| \geq 2^n/(n+1)$ and this gives the desired result.