Information Theory with Applications, Math6397
Lecture Notes from November 20th, 2014

taken by Nandini Deka

Last Time

- Linear codes for continuous channels.

Summary of the last result:

Theorem: Given a random input vector: $X : \Omega \to \mathbb{R}^k$, X_j i.i.d. with zero mean , Gaussian components having covariance matrix: $C_X = \frac{S}{k}I$ and a channel $\gamma : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$, $\gamma(\hat{X}) = \hat{X} + N$ then the best linear encoding achieves :

$$C(S) = \sum_{j=1}^{k} \frac{1}{2} ln\Big((\frac{1}{\mu\sigma_j^2} - \frac{k}{S})^+ + 1 \Big)$$

where $\{\sigma_j^2\}_{j=1}^k$ are the k smallest eigen values of C_N and μ is chosen such that :

$$\sum_{j=1}^{k} \Big(\frac{1}{\mu k} - \frac{\sigma_j^2}{S} \Big)^+ = 1$$

where $\frac{1}{\sigma_j^2} = \lambda_j$

Note: If noise has components, $\sigma_j = \sigma$ then using a linear encoding into $\mathbb{R}^n$ with large 'n' is of no use.

8 Frames as Codes

Recall: A q-ary block code of length m, $\phi_n : \mathbb{A}^k \to \mathbb{A}^n$, $|A| = q$ has $rate$:

$$R = \frac{log_q(m)}{n} = \frac{log_q|A|^k}{n} = \frac{log_q(q)^k}{n} = \frac{k}{n}$$

(also known as '*coding rate*')
For invertible ϕ_n we have : $R \leq 1$ or $k \leq n$.

8.0.1 Definition:

If $\phi_n : \mathbb{F}^k \to \mathbb{F}^n$ and $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ then we say that ϕ_n has **dimensionless rate** $R = \frac{k}{n}$. If $\phi_n(a) = Ha$, where H is a $n \times k$ matrix, then we say the **code is linear** and H is called the **encoding matrix**.
In this case by linearity of H and Riesz representation theorem (for Hilbert spaces):

$$(Ha)_j =< a, f_j >$$

and if H is left-invertible, then there exists $c > 0$ such that: $H^*H \geq cI$ where H is a positive definite matrix and we have an operator inequality, i.e.

$$< H^*Hx, x >= ||Hx||^2 \geq c||x||^2$$

where $\sqrt{c}$ is the strictly positive distance of x from 0.
Also by boundedness of H , there is $C > 0$, such that : $H^*H \leq CI$. Thus for each x in $\mathbb{F}^k$:

$$c||x||^2 \leq ||Hx||^2 = \sum_{j=1}^{n} | < x, f_j > |^2 \leq C||x||^2$$

If these inequalities hold for $0 < c \leq C < \infty$, then we say that $\{f_j\}_{j=1}^n$ forms a **frame for** $\mathbb{F}^k$.

Remark: Given a frame $\{f_j\}_{j=1}^n$ for $\mathbb{F}^k$, then : $\sum_{j=1}^n < x, f_j > f_j = H^*Hx$, where H^*H is an invertible matrix. So we have,for all x in $\mathbb{F}^k$:

$$\sum_{j=1}^{n} < x, f_j > (H^*H)^{-1} f_j = x$$

8.0.2 Definition:

For $\{f_j\}_{j=1}^n$ frame for $\mathbb{F}^k$, with associated map : $H : \mathbb{F}^k \to \mathbb{F}^n$ we call : $g_j = (H^*H)^{-1} f_j$ the **canonical dual frame** to $\{f_j\}_{j=1}^n$.

We can restore the signal with with a different set of vectors. For example, if K is an $n \times n$-matrix such that $HH^*K = 0$, then

$$\sum_{j,l=1}^{n} (\delta_{j,l} + K_{j,l}) < x, f_l > (H^*H)^{-1} f_j = x$$

and thus $h_j = \sum_{l=1}^n (\delta_{l,j} + K_{l,j})(H^*H)^{-1} f_l$ is another choice for a dual.

8.0.3 Definition:

If a frame is identical to its canonical dual, $g_j = f_j$ for all j, then we call $\{f_j\}_{j=1}^n$ a **Parseval frame**.
In this case, using $H^*H = I_k$ gives:

$$\sum_{j=1}^{n} < x, f_j >< f_j, x >=< x, x >= ||x||^2$$

similar to the Parseval identity for orthonormal basis, though the frame vectors do not have to be orthonormal.

8.0.4 Lemma

If $\{f_j\}_{j=1}^n$ is a Parseval frame for $\mathbb{F}^k$ and H the associated encoding matrix, then among all left inverses of H, H^* has minimal operator norm and Hilbert-Schmidt norm.

Proof: Let G ($k \times n$ matrix) be another left-inverse of H or $GH = I_k$.

- For the **Operator Norm:**

$$||G|| = \max_{\substack{y \in \mathbb{F}^n \\ y \neq 0}} \frac{||Gy||}{||y||} \geq \max_{\substack{y \in \mathbb{F}^n \\ y \neq 0 \\ y = Hx, x \in \mathbb{F}^k}} \frac{||Gy||}{||y||} \quad \text{(since we maximize over a lesser range now)}$$

$$= \max_{\substack{x \in \mathbb{F}^k \\ x \neq 0}} \frac{||GHx||}{||Hx||} \quad \text{(Here } GH = I \text{ and } ||Hx|| = ||x|| \text{ by Parseval)}$$

$$= \max_{\substack{x \in \mathbb{F}^k \\ x \neq 0}} \frac{||x||}{||x||}$$

or $||G|| \geq 1$

Also : $||H^*|| = ||H|| = 1$
Thus we have: $||G|| \geq ||H^*||$, or H^* has the minimal operator norm.

- For the **Hilbert-Schmidt Norm:** For any orthonormal basis $\{e_i\}_{i=1}^k$ of $\mathbb{F}^k$, using that $\{He_i\}_{i=1}^k$ is an orthonormal system (not basis), we get:

$$tr[G^*G] \geq \sum_{i=1}^{k} < G^*GHe_i, He_i >$$

$$= \sum_{i=1}^{k} < GHe_i, GHe_i > \quad \text{(using } GH = I\text{)}$$

$$= \sum_{i=1}^{k} < e_i, e_i > \quad (< e_i, e_i >= 1 \text{ since orthonormal})$$

$$= k = tr[H^*H]$$

or $tr[G^*G] \geq tr[HH^*]$

Thus H^* is the optimal choice for Hilbert-Schmidt norm.

Parseval frames have an interesting geometric property.
Since the encoding matrix H is an isometry, $H^*H = I$, and we obtain the trace identity

$$k = tr[H^*H] = tr[HH^*] = \sum_{j=1}^{n} < f_j, f_j >= \sum_{j=1}^{n} ||f_j||^2 .$$

8.0.5 Corollary:

For a Parseval frame $\{f_j\}_{j=1}^n$, $\sum_{j=1}^n \|f_j\|^2 = k$.

Thus, we can think of $\{f_j\}_{j=1}^n$ as being a 'vector-valued' sphere.

8.0.6 Definition:

If a frame $\{f_j\}_{j=1}^n$ has only vectors of norm : $||f_j|| = c$, then we call it an **equal-norm frame**.

8.0.7 Corollary:

If $\{f_j\}_{j=1}^n$ is an equal-norm and Parseval frame, then $||f_j|| = \sqrt{\frac{k}{n}}$ for each j.

Proof: Since $\{f_j\}_{j=1}^n$ is an equal-norm frame, using $k = \sum_{j=1}^n ||f_j||^2$, we get:

$$k = n||f_j||^2$$
$$\Rightarrow ||f_j|| = \sqrt{\frac{k}{n}}$$

□

8.0.8 Definition:

We call a Parseval frame $\{f_j\}_{j=1}^n$ for $\mathbb{F}^k$ a **(n,k)-frame**.

8.1 Erasures

Given an ecoding with the Parseval frame by associated isometry, $V : \mathbb{R}^k \to \mathbb{R}^n$, consider the "*loss*" of frame co-efficients. This means, we cannot use the value of certain coefficients to reconstruct or even approximate a vector x. To model this in a nathematically concise form, we set the corresponding frame coefficients to zero.

8.1.1 Definition:

A **one erasure** E_l, indexed by $l \in \{1, 2, ..., n\}$ is a map $E_l : \mathbb{R}^k \to \mathbb{R}^n$ given by:

$$x \mapsto \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & 1 & & & & \vdots \\ \vdots & & & 0 & (l^{th} \text{ column}) & & \vdots \\ \vdots & & & & 1 & & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} x$$

A **general erasure** is a diagonal projection matrix.

Qs. Can we recover x from EVx ,where E is a general erasure and how?