Information Theory with Applications, Math6397
Lecture Notes from November 18, 2014

taken by Junyu Ding

We are getting close to the frame theory!

Recall from the last time we did in class:

♦ $\gamma(X) = \overbrace{HX}^{\text{matrix}} + \overbrace{N}^{\text{i.i.d., Gaussian with variance } \sigma^2}$

We had arrived at an upper bound $\prod_{j=1}^{k}(\underbrace{1+(\Delta)^2_{j,j}}_{\text{eigenvalues of } C_X: S_j^2} \underbrace{(W^t D^2 W)_{j,j}}_{\text{eigenvalues of } H^tH: \lambda_j})$

(Note: This estimate is achieved if C_X and H^tH are simultaneously diagonalizable.)

To compute the capacity, we want to maximize $\ln \prod_{j=1}^{k}(1+.........)$.

EulerLagrange equations give: $\lambda_i(1+S_j\lambda_j)^{-1} = \begin{cases} M, & S_j > 0 \\ \lambda_j, & else \end{cases}$,

so $S_j = \begin{cases} \frac{1}{\mu} - \frac{1}{\lambda_j}, & S_j > 0 \\ 0, & else \end{cases}$ and μ is choosen such that $\sum_{j=1}^{k} S_j = S$.

Now we compute the mutual information

$$\begin{aligned} I(X;Y) &= \frac{1}{2}\ln((2\pi e)^n \prod_{j=1}(\lambda_j(\mu^{-1}-\lambda_j^{-1})^+ + 1)) - \frac{1}{2}\ln((2\pi e)^n)) \\ &= \frac{1}{2}\ln(\prod_{j=1}^{k}\max\{1, \lambda_j\mu^{-1}\}) \\ &= \frac{1}{2}\sum_{j=1}^{k}(\ln(\lambda_j\mu^{-1}))^+ \end{aligned}$$

where μ gives $\sum_{j=1}^{k}(\frac{1}{\mu}-\frac{1}{\lambda_j})^+ = S$.

Now we get to today's inequality:

Given $A = A^t$, opt. on $\mathbb{H}$ and $P = P^*P$, $(P = P^2, P = P^*)$, $rankP = k$, We claim $tr[PA] \le \sum_{j=1}^{k} \lambda_j$, whose $\{\lambda_1, \lambda_2, \lambda_3, ..., \lambda_{k-1,\lambda}, \lambda_k\}$ are the k largest eigenvalues of A.

Proof: Let $A = \sum_{j=1}^{\infty} \lambda_j E_j$, without loss of generality $rankE_j = 1$, $E_j^* E_j = E_j$. If P is a rank-free projection, then $P = XX^*$, $|X| = 1$, and

$$
\begin{aligned}
\max_{P=xx^*} tr[PA] &= \max_{P=xx^*} \sum_{l=1}^{\infty} \underbrace{\langle PAe_l, e_l\rangle}_{\langle Ae_l, x\rangle \underbrace{\langle x, e_l\rangle}_{\star}} \\
&= \max_{\|x\|=1} \langle Ax, x\rangle \\
&= \max_{\|x\|=1} \sum_{j=1}^{\infty} \lambda_j \underbrace{\langle E_j x, x\rangle}_{\|E_j x\|^2 = P_j, \sum_{j=1}^{\infty} P_j = 1} \\
&\le \max_j \lambda_j.
\end{aligned}
$$

(Note: $\bigstar$ is achieved if we choose $e_1 = x, e_l \perp x, l \ge 2$)

If $rankP = k$, we have $tr[AP] = \sum_{j=1}^{\infty} \lambda_j \underbrace{tr[E_j p]}_{\le 1}$

and $\sum_{j=1}^{\infty} tr[E_j P] = tr[P] = k$, so we need to:

maximize $\sum_{j=1}^{\infty} \lambda_j P_j$, subject to $0 \le P_j \le 1, \sum_{j=1}^{\infty} P_j = k$.

In this case, we choose $P_j = 1$ when j belongs to the k largest eigenvalues, and else set $P_j = 0$, then $tr[PA] = \sum_{j=1}^{k} \lambda_j$.

7 Linear codes for parallel additive white noise channels

Assume we have input $X = (X_1, X_2, ..., X_k)$ with X_j i.i.d. zero-mean Gaussians and $C_X = \frac{S}{k} I$ and suppose $\gamma : \mathbb{R}^n \times \Omega \to \mathbb{R}^k$ is additive white Gaussian noise (AWGN) $\gamma(\hat{X}) = \hat{X} + N$. We have input constraint $tr[C_{\hat{X}}] \le S$ and noise coveriance C_N.

QUESTION: What is the best linear encoding $\hat{X} = HX^2$?

By power constraint and

$$tr[C_{\hat{X}}] = tr[H \underbrace{C_x}_{\frac{S}{k}I} H^t]$$
$$= \frac{S}{k} tr[HH^t] \leq S$$

We have $tr[HH^t] \leq k$.

According to the usual procedure, we then want to:

maximize $det(C_N + \frac{S}{k}HH^t)$, subject to $tr[HH^t] \leq k$.

Choosing a basis which diagonalizes C_N then gives:

$det(C_N + \frac{S}{k}HH^t) = det(D^2 + \frac{S}{k}\hat{H}\hat{H}^t)$ with $\widetilde{H} = DH$.

We assume eigenvalues of C_N that are strictly positive, otherwise the capacity would be infinite. In this case, we can factor them out of the determinant:

$$\begin{aligned}
det(C_N + \frac{S}{k}HH^t) &= det(D^2)det(I_n + \frac{S}{k}D^{-1}\widetilde{H}\widetilde{H}^t D^{-1}) \\
&= det(D^2)det(I_k + \frac{S}{k}\widetilde{H}^t D^{-2}\widetilde{H}) \\
&= det(D^2)det(I_k + \frac{S}{k}\widetilde{V}^t\widetilde{W}^t D^{-2}\widetilde{W}\widetilde{V}) \\
&\leq det(D^2)\prod_{j=1}^{k}(1 + \frac{S}{k}\widetilde{V}^2_{j,j}(\widetilde{W}^t D^{-2}\widetilde{W})_{j,j})
\end{aligned}$$

where $\widetilde{V}$ is diagonal positive, $\widetilde{W}$ is isometry, equality holds in hadamard's inequality if and only if $\widetilde{W}^t D^{-2}\widetilde{W}$ is diagonal.

In this case, the maximum value is achieved by picking the k largest eigenvalues of D^{-2}.

To achieve equality, we choose HH^t so that it is simultaneously diagonalizable with C_N and the range of H is the span of eigenvectors of C_N corresponding to the k largest eigenvalues of C_N^{-1}, meaning the k smallest eigenvalues of C_N.

Then, $\widetilde{H}^t\widetilde{H}$ is diagonal with entries such that the Euler-Lagrange equations are satisfied:

$\frac{S}{k}\lambda_j(1 + \frac{S}{k}(\widetilde{H}^t\widetilde{H})_{j,j}\lambda_j)^{-1} = \begin{cases} \mu, (\widetilde{H}^t\widetilde{H})_{j,j} > 0 \\ \\ \frac{S}{k}\lambda_j, else \end{cases}$,

so $\widetilde{H}^t\widetilde{H} = \begin{cases} \frac{1}{\mu} - \frac{k}{S\lambda_j}, (\widetilde{H}^t\widetilde{H})_{j,j} > 0 \\ \\ 0, else \end{cases}$.

We then get:

$$
\begin{aligned}
I(X;Y) &= \frac{1}{2}\ln((2\pi e)^n \prod_{j=1}^{k}(\lambda_j(\mu^{-1} - \frac{k}{S}\lambda_j^{-1})^+ + 1)\underbrace{det(D^2)}_{det(C_N)}) - \frac{1}{2}\ln((2\pi e)^n det(C_N)) \\
&= \frac{1}{2}\ln(\prod_{j=1}^{k}(\underbrace{\lambda_j(\mu^{-1} - \frac{k}{S}\lambda_j^{-1})^+}_{\frac{\lambda_j}{\mu} - \frac{k}{S}} + 1)) \\
&= \frac{1}{2}\sum_{j=1}^{k}\ln((\frac{\lambda_j}{\mu} - \frac{k}{S})^+ + 1)
\end{aligned}
$$

where μ is choosen such that $tr[\widetilde{H}^t\widetilde{H}] = \sum_{j=1}^{k}(\frac{1}{\mu} - \frac{k}{S\lambda_j})^+ = k$.