Recall from last time we discussed

- Capacity of AWGN
- Channel with correlated noise
- Treat eigenvalues of covariance matrix like variances in independent case

Warm up: Matrix Theory
We need to prove this statement:

5.7.32 Proposition
For any $n \times k$ matrix $T$ over $\mathbb{R}$ $\det (I_n + TT^t) = \det (I_k + T^tT)$.

Proof. Consider $\epsilon \geq 0$. We want to show $\det (I_n + \epsilon TT^t) = \det (I_k + \epsilon T^tT)$ so if $\epsilon = 0$ the statement is proven because the LHS=RHS=1.

Now consider $\epsilon > 0$. First we want to define the functions $f_n(\epsilon) = \det (I_n + \epsilon TT^t)$ and $f_k(\epsilon) = \det (I_k + \epsilon T^tT)$ then we will compute the derivative of $f_n$ at 0.

$$f_n'(0) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \det (I_n + \epsilon TT^t)$$

$$= \text{Tr}[TT^t]$$

$$= \text{Tr}[T^tT]$$ by elementary matrix production also

$$= \frac{d}{d\epsilon} \left|_{\epsilon=0} \det (I_k + \epsilon T^tT) \right.$$ which satisfies the same differential equation as $f_k$ at $t = 0$.

More generally, if we choose $A$ to be the inverse $A = (I_n + \epsilon TT^t)^{-1}$ then

$$f_n(\bar{\epsilon} + \epsilon) = \det(I_n + \epsilon TT^t + \bar{\epsilon}TT^t) = \det(A^{-1}) \det(I_n + A\epsilon TT^t)$$

and taking the derivative at $\epsilon = 0$ gives

$$f_n'(\bar{\epsilon}) = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \det(I_n + \epsilon TT^t + \bar{\epsilon}TT^t) = \det(I_n + \epsilon TT^t) \text{tr}[(I_n + \epsilon TT^t)^{-1} \epsilon TT^t]$$

We note that by equating terms in the power series expansion,

$$\text{tr}[(I_n + \epsilon TT^t)^{-1} \epsilon TT^t] = \text{tr}[(I_k + \epsilon T^tT)^{-1} \epsilon T^tT]$$
thus reversing the steps gives \( f_n'(\tilde{\varepsilon}) = f_n(\tilde{\varepsilon})g(\tilde{\varepsilon}) \) and \( f_k'(\tilde{\varepsilon}) = f_k(\tilde{\varepsilon})g(\tilde{\varepsilon}) \) with the same \( g \).

By the uniqueness of the solution to ordinary differential equations with Lipschitz-continuous expressions for the derivative, we see that \( f_k \) and \( f_n \) coincide. \( \square \)

5.7.33 Theorem

For a fixed \( n \times k \) matrix \( H \) (linear model coding) and \( \mathbb{R}^n \)-valued zero mean Gaussian Random Variable \( N \) with covariance

\[
\mathbb{E}[N_iN_j] = \delta_{i,j}
\]

The capacity of \( \gamma, \gamma(X) = HX + N \) subject to the input constraint \( \sum_{j=1}^{k} \mathbb{E}[X_j^2] \leq S \) is

\[
C(S) = \sum_{j=1}^{k} (\ln(\mu \lambda_j))^+
\]

where \( \{\lambda_j\}_{j=1}^{k} \) are eigenvalues of \( H^tH \) and \( \mu \) is chosen such that \( \sum_{j=1}^{k} (\mu - \lambda_j^{-1})^+ = S \).

Proof. Based on Mutual Information, as usual \( I(X;Y) = h(Y) - h(N) \). So we need to maximize \( h(Y) \). Using short hand notation \( X = (x_1, x_2, \ldots, x_k)^t \)

\( C_X = \mathbb{E}[XX^t] \) because \( (C_X)_{i,j} = \mathbb{E}[X_iX_j] \).

By the linearity of \( X \to HX \) we have

\[
(C_X)_{ij} = \mathbb{E}[(HX + N)_i(HX + N)_j] = \mathbb{E}[(HX)_i(HX)_j] + \delta_{i,j}
\]

\[
C_Y = HC_XH^t + I
\]

Thus, we have an upper bound for \( h(Y) \) by the differential entropy of a corresponding Gaussian with the same covariance matrix, \( h(Y) \leq \frac{1}{2}\ln((2\pi e)^n \det(I + HC_XH^t)) \) and this bound can be achieved by choosing \( X \) Gaussian, which implies that \( Y \) is Gaussian.

For a given \( H \) take SVD(Singular Value Decomposition) such that:

\[
H = ODW
\]

where \( W \) is isometry, \( O \) is orthogonal, \( D \) is diagonal and \( D \) contains entries \( \left\{ \sqrt{\lambda_i} \right\}^n \) where \( \lambda_i \) are non negative eigenvalues of \( HH^t = OD^2O^t \) where \( D^2 \) has diagonal matrix \( \lambda_i \) and 0 elsewhere.

Using the orthogonal invariance of determinants, we then have \( h(Y) = \frac{1}{2}\ln((2\pi e)^n \det(I + DWCDW^tD^2W)) \).

Next we factor \( C_X = VV^t \) by positive definiteness of \( C_X \). So reordering in the determinant as shown above results in \( \det(I_n + DWCDW^tD) = \det(I_n + DWVV^tD^tD) \)

\[
= \det(I_k + V^tW^tD^2W^tDV^tD).
\]

To continue, we choose an SVD of \( V \) such that \( V = O'\Delta O'' \) where \( \Delta \) is the diagonal matrix and again after using the invariance of determinants we apply Hadamard’s inequality, \( \det(I_k + V^tW^tD^2W) = \det(I_k + \Delta W^tD^2W) \)

\[
\leq \prod_{j=1}^{k} \left( 1 + \frac{(\Delta W^tD^2W\Delta)_{j,j}}{\Delta^2_{j,j}(W^tD^2W)_{j,j}} \right)
\]
So Equality holds in Hadamard’s Inequality if and only if $WD^2W^t$ is diagonal. To achieve equality we let $VV^t$ have same eigenbasis such as $W^tD^2W = H^tH$.

Next we wish to maximize

$$\prod_{j=1}^{k} \left( 1 + \frac{(\Delta^2)_{j,j}}{\text{eigenvalue of } C_X} \frac{(W^tD^2W)_{j,j}}{\text{eigenvalue of } H^tH} \right)$$

subject to

$$\sum_{j=1}^{k} (\Delta^2)_{j,j} = Tr[C_X] \leq S$$

The optimal choice of $C_X$ gives the Euler-Lagrange Equation:

$$\lambda_j(1 + (VV^t)_{j,j}\lambda_j)^{-1} = \begin{cases} \mu, & (VV^t)_{j,j} > 0 \\ \lambda_j, & \text{else} \end{cases}$$

So

$$(VV^t)_{j,j} = \begin{cases} \frac{1}{\mu} - \frac{1}{\lambda_j}, & (VV^t)_{j,j} > 0 \\ 0, & \text{else} \end{cases}$$

and $\mu$ is chosen such that $\sum_{j=1}^{k}(VV^t)_{j,j} = S$.

To end this the Hadamard’s inequality achieved if and only if the matrix is diagonal matrix. Thus, the upper bound is achieved if the matrix is diagonal with corresponding entries such that Euler Lagrange equations are satisfied. 

$\Box$