UNIVERSITY OF CANTERBURY
Dept. of Electrical and Computer Engineering

ENEL 673

Solutions to Assignment 3 **June, 2002**

1. Two Seemingly-Different MMSE Solutions

We have $\mathbf{y} = \mathbf{C}\cdot\mathbf{A}\cdot\mathbf{b} + \mathrm{n} = \mathbf{F}\cdot\mathbf{b} + \mathrm{n}$ where $\mathbf{C}$ is MxK and the noise covariance matrix is

$$\mathbf{R}_{\mathrm{n}} = \frac{1}{2}\cdot \mathrm{E}\left(\mathbf{n}^{\mathrm{H}}\cdot\mathbf{n}\right) = \mathbf{I}_{\mathbf{M}}$$

(a) First, estimate the data bit b_1 from $\mathbf{y}$. Form $\mathrm{bhat}_1 = \overline{\mathbf{w_1}^{\mathrm{T}}}\cdot\mathbf{y}$ The MMSE weight vector $\mathbf{w}_1$ is the solution of the MxM normal equations $\mathbf{R_y}\cdot\mathbf{w_1} = \mathbf{p_1}$

where $$\mathbf{R_y} = \frac{1}{2}\cdot \mathrm{E}_{\mathrm{b}\nu}\left(\mathbf{y}\cdot\mathbf{y}^{\mathrm{H}}\right) = \frac{1}{2}\cdot\mathbf{F}\cdot\mathbf{F}^{\mathrm{H}} + \mathbf{I}_{\mathbf{M}}$$

assuming that the b and ν ensembles are independent and that all bits are iid, and

$$\mathbf{p_1} = \frac{1}{2}\cdot \mathrm{E}_{\mathrm{b}\nu}\left(\mathbf{y}\cdot\overline{\mathrm{b}_1}\right) = \frac{1}{2}\cdot\mathbf{f_1}$$

To obtain all weight vectors at once, collect the columns of $\mathbf{W}$ and $\mathbf{P}$ to form the matrix normal equation

$$\mathbf{R_y}\cdot\mathbf{W} = \mathbf{P} \qquad \text{or} \qquad \mathbf{W} = \left(\mathbf{F}\cdot\mathbf{F}^{\mathrm{H}} + 2\cdot\mathbf{I}_{\mathbf{M}}\right)^{-1}\cdot\mathbf{F}$$

The estimate is therefore

$$\mathbf{bhat} = \mathbf{F}^{\mathrm{H}}\cdot\left(\mathbf{F}\cdot\mathbf{F}^{\mathrm{H}} + 2\cdot\mathbf{I}_{\mathbf{M}}\right)^{-1}\cdot\mathbf{y} \qquad (1)$$

Next, estimate the data from the vector of sufficient stats $\mathrm{z} = \mathbf{F}^{\mathrm{H}}\cdot\mathbf{y} = \mathbf{F}^{\mathrm{H}}\cdot\mathbf{F}\cdot\mathbf{b} + \mathbf{F}^{\mathrm{H}}\cdot\mathrm{n}$

The normal equations follow the same pattern as above. We need the Gram matrix

$$\mathbf{R}_{\mathrm{z}} = \frac{1}{2}\cdot \mathrm{E}_{\mathrm{b}\nu}\left(\mathrm{z}\cdot\mathrm{z}^{\mathrm{H}}\right) = \frac{1}{2}\cdot\mathbf{F}^{\mathrm{H}}\cdot\mathbf{F}\cdot\mathbf{F}^{\mathrm{H}}\cdot\mathbf{F} + \mathbf{F}^{\mathrm{H}}\cdot\mathbf{F}$$

and the cross correlation matrix

$$\mathbf{P_z} = \frac{1}{2}\cdot E_{b\nu}\left(\mathbf{y}\cdot\mathbf{b}^H\right) = \frac{1}{2}\cdot\mathbf{F}^H\cdot\mathbf{F}$$

Then we have the normal equations

$$\mathbf{R_z}\cdot\mathbf{W_z} = \mathbf{P_z} \qquad \text{or} \qquad \left(\frac{1}{2}\cdot\mathbf{F}^H\cdot\mathbf{F}\cdot\mathbf{F}^H\cdot\mathbf{F} + \mathbf{F}^H\cdot\mathbf{F}\right)\cdot\mathbf{W_z} = \frac{1}{2}\cdot\mathbf{F}^H\cdot\mathbf{F}$$

With probability one, $\frac{1}{2}\cdot\mathbf{F}^H\cdot\mathbf{F}$ is non-singular, given its statistics, so we can multiply both sides of the equation by its inverse, to obtain the *K*x*K* set

$$\left(\mathbf{F}^H\cdot\mathbf{F} + 2\cdot\mathbf{I_K}\right)\cdot\mathbf{W_z} = \mathbf{I_K} \qquad \text{or} \qquad \mathbf{W_z} = \left(\mathbf{F}^H\cdot\mathbf{F} + 2\cdot\mathbf{I_K}\right)^{-1}$$

Then the estimate is

$$\mathbf{bhat} = \mathbf{W_z}\cdot z = \left(\mathbf{F}^H\cdot\mathbf{F} + 2\cdot\mathbf{I_K}\right)^{-1}\cdot\mathbf{F}^H\cdot\mathbf{y} \tag{2}$$

In (1) and (2), we have alternative forms of MMSE estimates of **b**. There is less computation in (2), since the matrix to be inverted is smaller. But do they produce the same estimate? That's what we determine in part (b).

(b) We can demonstrate that (1) and (2) produce the same result by expanding **F** in a singular value decomposition. For the following discussion, assume that $K \le M$, although it still holds if $K > M$. From the SVD, we have

$$\mathbf{F} = \mathbf{U}\cdot S\cdot\mathbf{V}^H = \begin{pmatrix}\mathbf{U_s} & \mathbf{U_0}\end{pmatrix}\cdot\begin{pmatrix}S_{\mathbf{s}} \\ \mathbf{0}\end{pmatrix}\cdot\mathbf{V}^H = \mathbf{U_s}\cdot S_{\mathbf{s}}\cdot\mathbf{V}^H \tag{3}$$

where

* **V** is *K*x*K* with orthonormal columns that span the input (**b**) space.

* S is *M*x*K* and contains the *K*x*K* submatrix S_s, a diagonal matrix of the *K* real non-negative singular values (actually, positive real with probability one for our problem) $\sigma_1, \sigma_2, \ldots \sigma_K$.

* **U** is *M*x*M* with orthonormal columns, of which the first *K* form the *M*x*K* submatrix $\mathbf{U}_s$ and the remainder form $\mathbf{U}_0$. The columns of **U** span the output (**y**) space, and the columns of $\mathbf{U}_s$ span the signal subspace (the space of images **Fb** of the vectors in the input space).

We will substitute the SVD of **F** into (1) and (2) to see if they are the same.

It's easier to start with (2), so make the substitution, using the last equality of (3). The weight matrix becomes

$$\mathbf{bhat} = \left(\mathbf{F}^H\cdot\mathbf{F} + 2\cdot\mathbf{I_K}\right)^{-1}\cdot\mathbf{F}^H\cdot\mathbf{y} = \left(\mathbf{V}\cdot S_{\mathbf{s}}{}^H\cdot\mathbf{U_s}{}^H\cdot\mathbf{U_s}\cdot S_{\mathbf{s}}\cdot\mathbf{V}^H + 2\cdot\mathbf{I_K}\right)^{-1}\cdot\mathbf{V}\cdot S_{\mathbf{s}}{}^H\cdot\mathbf{U_s}{}^H\cdot\mathbf{y}$$

$$\blacksquare = \left(\mathbf{V} \cdot S_{\mathbf{s}}{}^{2} \cdot \mathbf{V}^{H} + 2 \cdot \mathbf{I_K}\right)^{-1} \cdot \mathbf{V} \cdot S_{\mathbf{s}} \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y}$$

since $\mathbf{U}_s$ is unitary and S_s is real and diagonal

$$\blacksquare = \left(\mathbf{V} \cdot S_{\mathbf{s}}{}^{2} \cdot \mathbf{V}^{H} + 2 \cdot \mathbf{V} \cdot \mathbf{I_K} \cdot \mathbf{V}^{H}\right)^{-1} \cdot \mathbf{V} \cdot S_{\mathbf{s}} \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y}$$

since **V** is unitary

$$\blacksquare = \mathbf{V} \cdot \left(S_{\mathbf{s}}{}^{2} + 2 \cdot \mathbf{I_K}\right)^{-1} \cdot \mathbf{V}^{H} \cdot \mathbf{V} \cdot S_{\mathbf{s}} \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y}$$

since **V** is unitary

$$\blacksquare = \mathbf{V} \cdot \left(S_{\mathbf{s}}{}^{2} + 2 \cdot \mathbf{I_K}\right)^{-1} \cdot S_{\mathbf{s}} \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y}$$

since **V** is unitary

$$\blacksquare = \mathbf{V} \cdot \mathrm{diag}\left(\frac{\sigma_1}{\sigma_1{}^{2} + 2}, \blacksquare, \blacksquare, \blacksquare, \frac{\sigma_K}{\sigma_K{}^{2} + 2}\right) \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y} \qquad (4)$$

since $S_{\mathbf{s}}$ and $\mathbf{I_K}$ are diagonal, making the inverse matrix simple. This has a nice intepretation. Since the vectors in $\mathbf{U_s}$ are an orthonormal basis of the signal subspace, the operation $\mathbf{U_s}{}^{H} \cdot \mathbf{y}$ obtains the components of the projection of **y** onto the signal subspace. Each component is then scaled by a function of its singular value (the scale factor, or gain, of that one-dimensional subspace) - in effect, a MMSE estimate is applied to each component separately in this basis. Finally, the MMSE estimates multiply the columns of **V** to produce the estimate **bhat** in the input space.

Now, we try to do the same for equation (1). Substitute the first equality of (3) into (1):

$$\mathbf{bhat} = \mathbf{F}^{H} \cdot \left(\mathbf{F} \cdot \mathbf{F}^{H} + 2 \cdot \mathbf{I_M}\right)^{-1} \cdot \mathbf{y}$$

$$\blacksquare = \mathbf{V} \cdot S^{H} \cdot \mathbf{U}^{H} \cdot \left(\mathbf{U} \cdot S \cdot \mathbf{V}^{H} \cdot \mathbf{V} \cdot S^{H} \cdot \mathbf{U}^{H} + 2 \cdot \mathbf{I_M}\right)^{-1} \cdot \mathbf{y}$$

$$\blacksquare = \mathbf{V} \cdot S^{H} \cdot \mathbf{U}^{H} \cdot \left(\mathbf{U} \cdot S^{2} \cdot \mathbf{U}^{H} + 2 \cdot \mathbf{I_M}\right)^{-1} \cdot \mathbf{y}$$

$$\blacksquare = \mathbf{V} \cdot S^{H} \cdot \left(S^{2} + 2 \cdot \mathbf{I_M}\right)^{-1} \cdot \mathbf{U}^{H} \cdot \mathbf{y}$$

since **U** is unitary

$$\blacksquare = \mathbf{V} \cdot \left(S_{\mathbf{s}} \quad \mathbf{0}\right) \cdot \left(S^{2} + 2 \cdot \mathbf{I_M}\right)^{-1} \cdot \mathbf{U}^{H} \cdot \mathbf{y}$$

and note that S^2 is diagonal

$$\blacksquare = \mathbf{V} \cdot \mathrm{diag}\left(\frac{\sigma_1}{\sigma_1{}^{2} + 2}, \blacksquare, \blacksquare, \blacksquare, \frac{\sigma_K}{\sigma_K{}^{2} + 2}\right) \cdot \mathbf{U_s}{}^{H} \cdot \mathbf{y} \qquad (5)$$

This is the same expression as (4), so the two formulations of the MMSE estimate produce the same result.

2. ZF MUD With Imperfect CSI

We have the estimate (use subscript *h* to denote estimate "hat")

$$\mathbf{b_h} = \mathbf{A}^{-1} \cdot \left(\mathbf{C_h}^{H} \cdot \mathbf{C_h} \right)^{-1} \cdot \mathbf{C_h}^{H} \cdot \mathbf{y}$$

$$\blacksquare = \mathbf{A}^{-1} \cdot \left(\mathbf{C_h}^{H} \cdot \mathbf{C_h} \right)^{-1} \cdot \mathbf{C_h}^{H} \cdot \left(\mathbf{C_h} \cdot \mathbf{A} \cdot \mathbf{b} + \mathtt{m} + \mathtt{n} \right) \qquad \text{where} \qquad \mathtt{m} = \mathbf{E} \cdot \mathbf{A} \cdot \mathbf{b}$$

$$\blacksquare = \mathbf{b} + \mathtt{a} + \mathtt{b}$$

where $\mathtt{a} = \mathbf{A}^{-1} \cdot \left(\mathbf{C_h}^{H} \cdot \mathbf{C_h} \right)^{-1} \cdot \mathbf{C_h}^{H} \cdot \mathtt{m} \qquad \mathtt{b} = \mathbf{A}^{-1} \cdot \left(\mathbf{C_h}^{H} \cdot \mathbf{C_h} \right)^{-1} \cdot \mathbf{C_h}^{H} \cdot \mathtt{n}$

Channel estimation error has produced $\mathtt{m}$, another additive disturbance. Is it noise-like? What are its statistics? First, it is at least conditionally Gaussian (conditioned on **b**). Closer inspection shows that the random phases in **E** make $\mathtt{m}$ independent of **b**, which contains only +/-1 components, so it is unconditionally Gaussian. Second, it has zero mean, since **E** has zero mean. Finally, its covariance matrix is

$$\mathbf{R}_{\mathtt{m}} = \frac{1}{2} \cdot E_{eb} \left(\mathbf{E} \cdot \mathbf{A} \cdot \mathbf{b} \cdot \mathbf{b}^{H} \cdot \mathbf{A}^{H} \cdot \mathbf{E}^{H} \right) = \frac{1}{2} \cdot E_{e} \left(\mathbf{E} \cdot \mathbf{A}^{2} \cdot \mathbf{E}^{H} \right) \qquad \text{since the data components are i.i.d.}$$

The *i,k* component of this matrix is just

$$\mathbf{R}_{\mathtt{m}_{i,k}} = \frac{1}{2} \cdot E_{e} \left(\sum_{j} A_j^{2} \cdot e_{i,j} \cdot \overline{e_{k,j}} \right) = \sigma_e^{2} \cdot \sum_{j} A_j^{2} \cdot \delta_{i,k}$$

so $\mathbf{R}_{\mathtt{m}} = \sigma_e^{2} \cdot \sum_{j} A_j^{2} \cdot \mathbf{I}$. We have equipower signals, so $\mathbf{A} = \sqrt{2 \cdot E_s} \cdot \mathbf{I}$, which we can substitute to obtain $\mathbf{R}_{\mathtt{m}} = \sigma_e^{2} \cdot K \cdot 2 \cdot E_s \cdot \mathbf{I}$

So the new noise μ acts just like the AWGN ν, just with a different variance. Effectively, we have done no more than reduce the SNR.

What is the effective SNR, compared with the original SNR? The original noise variance was unity, so the new noise variance is $1 + 2 \cdot K \cdot \sigma_e^{2} \cdot E_s$ and the new SNR is

$$\Gamma'_s = \frac{E_s}{\left(1 + 2 \cdot K \cdot \sigma_e^{2} \cdot E_s \right)} = \frac{\Gamma_s}{\left(1 + 2 \cdot K \cdot \sigma_e^{2} \cdot \Gamma_s \right)}$$

where the second equality uses the fact that E_S equals the original SNR, since the original noise variance was unity. This looks right. For zero estimation error variance, it reduces to the original SNR Γ_S. If it is non-zero, then even pumping the original SNR up to huge values leaves the altered SNR at $1/2K\sigma_e^2$ - that is, there is an error floor.

(b) As for the BER, the white noise equivalence of μ means that our previous expressions for BER of ZF MUD hold, although with an altered SNR. They were the same as single-user maximal ratio combining with M-K+1 degrees of freedom), so that

$$\mu(\Gamma) = \sqrt{\frac{\Gamma'_s}{1+\Gamma'_s}} \qquad \mathrm{BC}(n,k) = \frac{n!}{k!\cdot(n-k)!}$$

$$P_{err}(\Gamma',M) = \left(\frac{1-\mu(\Gamma')}{2}\right)^M \cdot \sum_{m=0}^{M-K} \mathrm{BC}(M-1+m,m)\cdot\left(\frac{1+\mu(\Gamma')}{2}\right)^m$$

(c) There are many ways to answer this question. Here is one. For any target BER, such as 10^{-3}, there is a corresponding target SNR Γ_t for a ZF system with a certain number on antennas and perfect channel state information. The modified SNR must at least equal that figure, so we obtain a relation between Γ_s and σ_e^2,

$$\Gamma_t \le \frac{\Gamma_s}{\left(1+2\cdot K\cdot\sigma_e^2\cdot\Gamma_s\right)}$$

As σ_e^2 increases from zero, Γ_s must increase to compensate. It doesn't work forever, though - the best we can do is $\Gamma_s \to \infty$, so that

$$\Gamma_t \le \frac{1}{2\cdot K\cdot\sigma_e^2} \qquad \text{or} \qquad \sigma_e^2 \le \frac{1}{2\cdot K\cdot\Gamma_t}$$

To relate this to the required correlation coefficient ρ, we use

$$\sigma_e^2 = \left(1-\rho^2\right)\cdot\sigma_c^2$$

Since $\sigma_c^2 = 1/2$, we have the absolute minimum requirement

$$\rho \ge \sqrt{1-\left(K\cdot\Gamma_t\right)^{-1}} = 1 - \frac{1}{2\cdot K\cdot\Gamma_t} \qquad \text{(approx)}$$

As an example, for M=2 antennas, we need about 13 dB SNR for 10^{-3}. This is 20 in natural units, so that ρ cannot drop below 0.988 - and it should be considerably better than that, since we don't want to use infinite transmit power.

(d) From (c), we see that the variance of the additive disturbance m is proportional to the product of estimation error variance and SNR, at least for equipower signals. From this, we expect that stronger signals generate greater contributions to additive noise. Going back to (a), we can make this quantitative. The covariance matrix is

$$\mathbf{R}_m = \sigma_e^2 \cdot \sum_j A_j^2 \cdot \mathbf{I}$$

This is white noise, as before, so the variance of the estimation error disturbance falls equally and randomly on the two bits. This means that the two users experience the same effective noise variance, just as in the perfect CSI case, although that variance is greater here, given by

$$\sigma_e^2 \cdot \sum_j A_j^2 + 1 = 1 + 2 \cdot \sigma_e^2 \cdot \left(\Gamma_{strong} + \Gamma_{weak} \right)$$

The effective SNR values of the users differ, since one is strong and one is weak:

$$\Gamma'_{strong} = \frac{\Gamma_{strong}}{1 + 2 \cdot \sigma_e^2 \cdot \left(\Gamma_{strong} + \Gamma_{weak} \right)}$$

$$\Gamma'_{strong} = \frac{\Gamma_{weak}}{1 + 2 \cdot \sigma_e^2 \cdot \left(\Gamma_{strong} + \Gamma_{weak} \right)}$$

The weaker user, being weaker, has a much poorer error rate. So, in summary,

* The stronger user experiences an effective estimation error noise level that is almost proportional to its own power, with a small contribution from the weaker user. As the power of both users increases, the Gaussian noise contribution becomes negligible, by comparison, and we are into an error floor region.

* The weaker user experiences the the same effective estimation error noise level, which is proportional to the power of the stronger user. Too bad for the weaker user. With perfect CSI, it might have had a good SNR. With channel estimation error, depending on the correlation coefficient ρ, its SNR might be very low indeed.

A complicating factor, which I didn't address, is the fact that the ρ for the stronger user is typically much better than the ρ for the weaker user.