SIMON FRASER UNIVERSITY
School of Engineering Science

ENSC 428 Data Communications

Solutions to Assignment 2 **February 2001**

1. Sum of Correlated Random Variables

(a) Denote the sample spacing by t_s, so $\mathbf{x} = \left(x(t_s) \;\; x(2\cdot t_s) \;\; \cdots \;\; x(N\cdot t_s) \right)^T$

For N=5, its covariance matrix is

$$\mathbf{C} = E\left(\mathbf{x}\cdot\mathbf{x}^T\right) = \begin{pmatrix} R_X(0) & R_X(t_s) & R_X(2\cdot t_s) & R_X(3\cdot t_s) & R_X(4\cdot t_s) \\ R_X(-t_s) & R_X(0) & R_X(t_s) & R_X(2\cdot t_s) & R_X(3\cdot t_s) \\ R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) & R_X(t_s) & R_X(2\cdot t_s) \\ R_X(-3\cdot t_s) & R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) & R_X(t_s) \\ R_X(-4\cdot t_s) & R_X(-3\cdot t_s) & R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) \end{pmatrix}$$

Matrices like this, in which all the elements of a diagonal are the same, are termed *Toeplitz*. Ours also symmetric, a combination that has interesting and useful properties. In any case, if the autocorrelation function is negligible after, say, two samples, then we have only five non-zero diagonals, no matter how large N becomes:

$$\mathbf{C} = \begin{pmatrix} R_X(0) & R_X(t_s) & R_X(2\cdot t_s) & 0 & 0 & 0 \\ R_X(-t_s) & R_X(0) & R_X(t_s) & R_X(2\cdot t_s) & 0 & 0 \\ R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) & R_X(t_s) & R_X(2\cdot t_s) & 0 \\ 0 & R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) & R_X(t_s) & R_X(2\cdot t_s) \\ 0 & 0 & R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) & R_X(t_s) \\ 0 & 0 & 0 & R_X(-2\cdot t_s) & R_X(-t_s) & R_X(0) \end{pmatrix}$$

(b) If $y = \mathbf{e}^T \cdot \mathbf{x}$ then the variance of *y* is

$$\sigma_y{}^2 = E\left(\mathbf{e}^T \cdot \mathbf{x} \cdot \mathbf{x}^T \cdot \mathbf{e}\right) = \mathbf{e}^T \cdot \mathbf{C} \cdot \mathbf{e} = \sum_{i=1}^{N} \sum_{k=1}^{N} C_{i,k}$$

If only the autocorrelation function is non-negligible only up to *d* samples, then

$$\sigma_y{}^2 = N \cdot R_x(0) + (N-1) \cdot R_x(t_s) + (N-1) \cdot R_x(-t_s) \ldots$$
$$+ (N-d) \cdot R_x(d \cdot t_s) + (N-d) \cdot R_x(-d \cdot t_s)$$

When *N* becomes very large, then

$$\sigma_y{}^2 = N \cdot \left(R_x(-d \cdot t_s) + \blacksquare + R_x(-t_s) + R_x(0) + R_x(t_s) + \blacksquare + R_x(d \cdot t_s)\right) \qquad \text{(approx)}$$

which is proportional to *N*.

2. Maximization of SNR

(a) To make it concrete, use *a*=1. Then the mean value i: $m_d = \sum_{i=1}^{N} w_i \cdot m_i$ and

$$\gamma = \frac{\left(\sum_{i=1}^{N} w_i \cdot m_i\right)^2}{\sigma^2 \cdot \sum_{i=1}^{N} (w_i)^2}$$

Note that scaling all the weights by a common factor scales both the squared mean (the signal power) and the noise variance by the square of that factor but leaves γ unchanged. We can theref maximize γ by maximizing its numerator with a constraint on its denominator, which we might as set to σ^2. Using a Lagrange multiplier λ, we maximize

$$J = \left(\sum_{i=1}^{N} w_i \cdot m_i\right)^2 - \lambda \cdot \sigma^2 \left[\sum_{i=1}^{N} (w_i)^2 - 1\right]$$

There are a few ways to get at this. Here's a concise way, but I'll do the expanded version furt below. Denote the column vectors of weights and means as **w** and **m,** respectively. Then

$$J = \mathbf{w}^T \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w} - \lambda \cdot \sigma^2 \cdot \left(\mathbf{w}^T \cdot \mathbf{w} - 1\right)$$

Setting the gradient with respect to **w** to zero gives

$$\nabla \mathbf{J} = 2 \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w} - 2 \cdot \lambda \cdot \sigma^2 \cdot \mathbf{w} = \mathbf{0}$$ (note this is a zero vector)

That is, we obtain an eigenvalue problem

$$\mathbf{M} \cdot \mathbf{w} = \lambda \cdot \sigma^2 \cdot \mathbf{w} \qquad \text{where} \qquad \mathbf{M} = \mathbf{m} \cdot \mathbf{m}^T$$

The matrix **M** has rank equal to 1, since any vector **w** orthogonal to **m** produces zero when pre-multiplied by **M**, and we can find *N*-1 linearly independent such vectors. Thus there are *N*-1 eigenvalues of **M** that equal zero. There is only one non-zero eigenvalue, and its eigenvector is proportional to **m**. So make **w** proportional to **m**. More simply, just observe that setting **w**=*c***m**

$$c \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{m} = c \cdot \lambda \cdot \sigma^2 \cdot \mathbf{m} \qquad \text{which we can satisfy with } \lambda = \frac{\mathbf{m}^T \cdot \mathbf{m}}{\sigma^2}$$

Substituting this **w** into our definition of SNR γ gives us

$$\gamma_{max} = \frac{\left(\mathbf{m}^T \cdot \mathbf{m}\right)^2}{\sigma^2 \cdot \mathbf{m}^T \cdot \mathbf{m}} = \frac{\mathbf{m}^T \cdot \mathbf{m}}{\sigma^2}$$

(improves with increasing number of measurements *N*, provided means m_i don't go to zero)

Now back to the expanded version of the problem. Using the definition of *J* on the previous p form each of the partial derivatives in turn (this is equivalent to the gradient, of course):

$$\frac{d}{dw_1} J = 2 \cdot \left(\sum_{i=1}^{N} w_i \cdot m_i \right) \cdot m_1 - 2 \cdot \lambda \cdot \sigma^2 w_1 = 0$$

$$\frac{d}{dw_2} J = 2 \cdot \left(\sum_{i=1}^{N} w_i \cdot m_i \right) \cdot m_2 - 2 \cdot \lambda \cdot \sigma^2 w_2 = 0$$ and so on.

This is a set of linear equation in the w_i, so collect terms and write them neatly

$$\left[(m_1)^2 - \lambda\cdot\sigma^2\right]\cdot w_1 + m_2\cdot m_1\cdot w_2 + \blacksquare + \blacksquare + m_N\cdot m_1\cdot w_N = 0$$

$$m_1\cdot m_2\cdot w_1 + \left[(m_2)^2 - \lambda\cdot\sigma^2\right]\cdot w_2 + \blacksquare + \blacksquare + m_N\cdot m21\cdot w_N = 0$$

$$\blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare$$

$$m_1\cdot m_N\cdot w_1 + m_2\cdot m_N\cdot w_2 + \blacksquare + \blacksquare + \left[(m_N)^2 - \lambda\cdot\sigma^2\right]\cdot w_N = 0$$

or in matrix form

$$\left(\mathbf{M} - \lambda\cdot\sigma^2\cdot\mathbf{I}\right)\cdot\mathbf{w} = 0 \qquad \text{where} \qquad \mathbf{M} = \mathbf{m}\cdot\mathbf{m}^T$$

Again, recognize this as an eigenvalue problem and reason it through the same way to obtain the weights as

$$w_i = m_i \quad \text{for} \quad i = 1..N$$

We now know that the weights should be proportional to the mean values. This is a very important result, and it sets the scene for matched filters, antenna arrays and many other statistica problems involving maximization of SNR. **For a solution in which the variables are complex, instead of real, see the Appendix .**

(b) If the noise variances are not all the same, the problem is a little tougher, but the principles ar same. We form

$$d = \sum_{i=1}^{N} w_i\cdot x_i$$

We can transform it back to the problem in part (a) simply by scaling each of the variables; doing loses no information, since the signal and noise in each variable are scaled by the same amount. I

$$v_i = \frac{x_i}{\sigma_i} \qquad \text{which has unit variance and mean} \qquad b_i = \frac{m_i}{\sigma_i}$$

From the results of part (a), we would multiply the v_i variables by weights proportional to b_i. Tha equivalent to multiplying the original x_i variables by weights

$$w_i = \frac{m_i}{(\sigma_i)^2}$$

We can obtain the same result with matrix notation. The SNR is now

$$\gamma = \frac{\left(\sum_{i=1}^{N} w_i \cdot m_i\right)^2}{\sum_{i=1}^{N} (w_i)^2 \cdot (\sigma_i)^2} = \frac{\mathbf{w}^T \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w}}{\mathbf{w}^T \cdot \mathbf{S} \cdot \mathbf{w}} \qquad \text{where} \qquad \mathbf{S} = \mathrm{diag}\left(\sigma_1{}^2, \blacksquare .. \blacksquare .. \sigma_N{}^2\right)$$

Using a Lagrange multiplier, we maximize

$$J = \mathbf{w}^T \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w} - \lambda \cdot \left(\mathbf{w}^T \cdot \mathbf{S} \cdot \mathbf{w} - 1\right) = \mathbf{w}^T \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w} - \lambda \cdot \left(\mathbf{w}^T \cdot \mathrm{S} \cdot \mathrm{S} \cdot \mathbf{w} - 1\right)$$

in which the noise covariance matrix is factored to become $\mathbf{S} = \mathrm{S}^2$, where

$$\mathrm{S} = \mathrm{diag}\left(\sigma_1, \sigma_2 .. \blacksquare .. \sigma_N\right)$$

Make a chage of variables $\quad \mathbf{u} = \mathrm{S} \cdot \mathbf{w} \quad \mathbf{w} = \mathrm{S}^{-1} \cdot \mathbf{u} \quad$ and rewrite *J* as

$$J = \mathbf{u}^T \cdot \mathrm{S}^{-1} \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathrm{S}^{-1} \cdot \mathbf{u} - \lambda \cdot \left(\mathbf{u}^T \cdot \mathbf{u} - 1\right) = \mathbf{u}^T \cdot \mathbf{b} \cdot \mathbf{b}^T \cdot \mathbf{u} - \lambda \cdot \left(\mathbf{u}^T \cdot \mathbf{u} - 1\right)$$

where $\quad \mathbf{b} = \mathrm{S}^{-1} \cdot \mathbf{m}$

Now it is just like the problem in part (a). The optimizing choice makes the transformed weig vector **u** proportional to **b**. Then

$$\mathbf{u_{opt}} = \mathrm{S}^{-1} \cdot \mathbf{m} \qquad \text{and transforming back,} \qquad \mathbf{w_{opt}} = \mathrm{S}^{-1} \cdot \mathbf{u_{opt}} = \mathbf{S}^{-1} \cdot \mathbf{m}$$

In other words, the optimum weight vector has coefficients

$$w_{opt_i} = \frac{m_i}{(\sigma_i)^2}$$

This is another very useful result, especially for coloured noise, time varying noise or different no
levels on several different measurements.

What is the resulting maximized SNR? Substitute this optimized choice for **w** back into the expression for γ and we get

$$\gamma = \frac{\mathbf{w_{opt}}^T \cdot \mathbf{m} \cdot \mathbf{m}^T \cdot \mathbf{w_{opt}}}{\mathbf{w_{opt}}^T \cdot \mathbf{S} \cdot \mathbf{w_{opt}}} = \frac{\left(\mathbf{m}^T \cdot \mathbf{S}^{-1} \cdot \mathbf{m}\right)^2}{\mathbf{m}^T \cdot \mathbf{S}^{-1} \cdot \mathbf{S} \cdot \mathbf{S}^{-1} \cdot \mathbf{m}} = \mathbf{m}^T \cdot \mathbf{S}^{-1} \cdot \mathbf{m}$$

Again, see the Appendix for a solution when the variables are complex.

3. Operation at Threshold

In PCM, operation at threshold is clearly attractive. If you are at threshold, consider the effec departing from it. First, if you are concerned with selecting the transmit power when the bandwic fixed, then increasing the power beyond the threshold point yields no improvement in output SNI On the other hand, if selecting bandwidth is the question when the power is fixed, then increasing bits per sample decreases the energy per bit and increases the BER, thereby degrading output SN but decreasing the number of bits gives more quantization noise. However, note that we don't alv operate at threshold; in some cases, threshold might involve a huge number of bits per sample, so are content to operate into the saturation region, provided that the number of bits is sufficient for output SNR.

In FM, it is slightly less clear, but again consider departures from threshold. If the transmit po is fixed, and you increase the modulation index, then the channel bandwidth and the total noise increases, causing below threshold operation and rapidly decreasing output SNR; yet decreasing t modulation index reduces the total noise in proportion but the differentiation gain by the square. way, you lose. On the other hand, if the modulation index (hence bandwidth) is fixed, and you increase the transmit power, you gain only a proportional increase in output SNR; not a kick in th head, but a little disappointing.

4. A Repeater Chain

To tackle this problem, I'll set up some definitions that are common to FM and PCM:

amp output signal power	noise PSD	signal bandwidth
$P_o := 10$ watt	$N_o := 2 \cdot 10^{-10}$ watt/Hz	$B := 10^6$ Hz

total length | # of repeaters | length of links

$D := 150$ km | M | $d(M) := \frac{D}{M}$

cable loss function

$L_o := 20$ dB/km $L(x) := 10^{0.1 \cdot L_o \cdot x}$ power loss, natural units, not dB

Transmission by FM

There are M amplifiers, each with a gain adjusted to compensate for the loss $L(x)$ in the precec cable section. The noise accumulates, and is proportional to the number of amplifiers traversed. overall noise PSD (at the input of the last amplifier) is $MN_o/2$.

The overall C/N at the final amp must be over 10 dB (note that this ratio is carrier power to nc power in the transmission bandwidth); that is,

$$\frac{C}{N} = \frac{P_o}{L(d(M)) \cdot N_o \cdot M \cdot B_c} \geq 10$$

Solve this for channel bandwidth as a function of the number of repeaters

$$B_c(M) := \frac{P_o}{L(d(M)) \cdot N_o \cdot M \cdot 10}$$

and for the modulation index (Carson's rule)

$$\beta(M) := \frac{B_c(M)}{2 \cdot B} - 1$$

Finally, the output SNR (assuming threshold and 3σ loading):

$$SNR_o = \frac{1}{3} \cdot \Gamma \cdot \beta^2 \qquad \text{where, at the last amp,} \qquad \Gamma = \frac{P_o}{L(d(M)) \cdot N_o \cdot M \cdot B}$$

Substituting this expression for Γ,

$$SNR_o = \frac{1}{3}\cdot\frac{P_o}{L(d(M))\cdot N_o\cdot M\cdot B}\cdot\frac{B_c}{B_c}\cdot\beta^2 = \frac{1}{3}\cdot 10\cdot\frac{B_c}{B}\cdot\beta^2 = \frac{10\cdot 2}{3}\cdot(\beta+1)\cdot\beta^2$$

and since β depends on the number of amps *M*:

$$SNR_o(M) := \frac{20}{3}\cdot(\beta(M)+1)\cdot\beta(M)^2$$

We can maximize this simply by maximizing β with respect to *M*. Referring to the expression for we see that we must maximize B_c, which we can do by minimizing its denominator; so find *M* that minimizes

$$M\cdot L(d(M)) = M\cdot 10^{0.1\cdot L_o\cdot\frac{D}{M}} = M\cdot\exp\left(0.1\cdot L_o\cdot\ln(10)\cdot\frac{D}{M}\right)$$

Treat *M* as continuous and set the derivative to zero:

$$\exp\left(0.1\cdot L_o\cdot\ln(10)\cdot\frac{D}{M}\right) - M\cdot\exp\left(0.1\cdot L_o\cdot\ln(10)\cdot\frac{D}{M}\right)\cdot 0.1\cdot L_o\cdot\ln(10)\cdot\frac{D}{M^2} = 0$$

which gives the optimum *M* as

$$M_{opt} := 0.1\cdot\ln(10)\cdot L_o\cdot D \qquad M_{opt} = 690.776$$

This is a lot of amps! They're every $d(M_{opt}) = 0.217$ km

Moreover, the output SNR is ridiculous: $SNR_o(M_{opt}) = 0.975$

Why did it turn out like this? Well, whenever you look for an optimum, you have to be clear of the phenomena involved. In our case, if we have too few amps, then the links are too long and the signal receive at the final amp is poor. If there are too many amps, then the accumulated noise is high. Check what happens as *M* varies (next page):

$M := 100, 200 .. 1000$

$M =$	$\beta(M) =$	$SNR_o(M) =$
100	-0.975	0.158
200	-0.605	0.964
300	-0.167	0.154
400	0.111	0.092
500	0.256	0.548
600	0.318	0.886
700	0.331	0.974
800	0.318	0.887
900	0.289	0.72
$1 \cdot 10^3$	0.253	0.535

Negative β for M<400?? It simply means that the signal is so weak that to stay over threshold requires a channel bandwidth lower than that of the signal itself. Clearly impossible.

Transmission by PCM

For any *M*, we get the distance, hence the SNR at the input to a regenerator. The trick is to obtain the corresponding number of bits so it operates at threshold (defined as a 1 dB drop). Eas approximate it with the overbound on the *Q* function.

First, get the SNR at each repeater $\Gamma(M) := \dfrac{P_o}{L(d(M)) \cdot N_o \cdot B}$

then approximate the Q function to give $P_e = Q\left(\sqrt{\dfrac{\Gamma}{n}}\right) = \dfrac{1}{2} \cdot \exp\left(\dfrac{-\Gamma}{2 \cdot n}\right)$ (approximately)

To apply the threshold condition, note that a 1 dB drop means that the transmission error term is of the quantization error term, so that

$$4 \cdot \frac{2^{2 \cdot n} - 1}{2^{2 \cdot n}} \cdot P_e = \frac{1}{4} \cdot 2^{-2 \cdot n}$$

At the end of our chain of *M* repeaters, approximate the overall BER as *M* times the BER on eacl link and the threshold condition becomes

$$2^{-2 \cdot n} = 16 \cdot M \cdot P_e = 8 \cdot M \cdot \exp\left(\frac{-\Gamma(M)}{2 \cdot n}\right)$$

Taking the logarithm and simplifying to form a quadratic equation in *n*, we obtain

$$n^2 + b \cdot n + c = 0$$

where

$$b(M) := \frac{\ln(8 \cdot M)}{2 \cdot \ln(2)} \qquad c(M) := \frac{-\Gamma(M)}{4 \cdot \ln(2)}$$

so the number of bits required for threshold operation with *M* repeaters is the solution of the quac

$$n(M) := \frac{-b(M) + \sqrt{b(M)^2 - 4 \cdot c(M)}}{2}$$ Note that only the positive root has meaning

Finally, the output SNR is simple if we assume operation at threshold:

$$SNR_o(M) := \frac{1}{3} \cdot \frac{1}{1.25 \cdot 2^{-2 \cdot n(M)}}$$ (recall 3σ loading)

In principle, there is an optimum value for *M*, as there was in FM, although establishing the fa requires tracing through some cumbersome limits as *M* becomes very large. However, that optim occurs with a huge number of repeaters, a ridiculously large number of bits per sample and a truly amazing output SNR. In practice, as we see in the tables below, we don't bother to approach this point, and we can safely work in the region where SNR_o increases with *M*. Some test values:

$$M := 100, 125 .. 300$$

M =	d(M) =	Γ(M) =	n(M) =	$SNR_o(M)$ =
100	1.5	50	2.472	8.212
125	1.2	199.054	6.34	$1.751 \cdot 10^3$
150	1	500	11.113	$1.308 \cdot 10^6$
175	0.857	965.349	16.229	$1.573 \cdot 10^9$
200	0.75	$1.581 \cdot 10^3$	21.367	$1.951 \cdot 10^{12}$
225	0.667	$2.321 \cdot 10^3$	26.354	$1.963 \cdot 10^{15}$
250	0.6	$3.155 \cdot 10^3$	31.102	$1.416 \cdot 10^{18}$
275	0.545	$4.056 \cdot 10^3$	35.571	$6.947 \cdot 10^{20}$
300	0.5	$5 \cdot 10^3$	39.752	$2.285 \cdot 10^{23}$

We see that 60 dB SNR_o is achievable with 150 amps (1 km spacing) and 11 or 12 bit linear quantization. A striking contrast to FM!

The reason why there is an optimum point, instead of a constant improvement? If there are to repeaters, then the signal level is too weak and the BER is too high. On the other hand, if the nur of repeaters is monstrously huge (with 1 centimetre spacing, say), then there is no appreciable cat loss, and the individual link BER is determined by the 10 watt received power. At this point, incr the number of amps simply increases the *accumulated* BER at the end of the chain, and the optim number of bits begins to decrease again - but it's a very shallow curve

APPENDIX TO QUESTION 2:
SNR MAXIMIZATION WHEN THE VARIABLES ARE COMPLEX

We form the inner product of the weight vector and the measurements as follows

$$d = \mathbf{w}^H \cdot \mathbf{x}$$

where the superscript *H* denotes conjugate transpose. The only difference from the earlier formu is the conjugate on the weights.

(a) If the variances are the same, we have

$$m_d = \mathbf{w}^H \cdot \mathbf{m} \qquad \sigma_d^{\,2} = \sigma^2 \cdot \sum_{i=1}^{N} \left(|w_i|\right)^2 = \sigma^2 \cdot \mathbf{w}^H \cdot \mathbf{w}$$

To maximize $\gamma = \dfrac{\left(|m_d|\right)^2}{\sigma_d^{\,2}}$ we maximize $J = \mathbf{w}^H \cdot \mathbf{m} \cdot \mathbf{m}^H \cdot \mathbf{w} - \lambda \cdot \sigma^2 \cdot \left(\mathbf{w}^H \cdot \mathbf{w} - 1\right)$

The gradient requires a little more care, since *J* is not analytic. You could differentiate with respe the real and imaginary parts separately, or use

$$\frac{d}{dw} w = 1 \qquad \frac{d}{d\overline{w}} \overline{w} = 1 \qquad \frac{d}{d\overline{w}} w = 0 \qquad \frac{d}{dw} \overline{w} = 0$$

where the overhead bar denotes conjugation in Mathcad. Either way, you get

$$Ñ\, \mathbf{J} = 2 \cdot \mathbf{m} \cdot \mathbf{m}^H \cdot \mathbf{w} - 2 \cdot \lambda \cdot \sigma^2 \cdot \mathbf{w} = \mathbf{0}$$

and the eigenvalue problem

$$\mathbf{M}\cdot\mathbf{w} = \lambda\cdot\sigma^2\cdot\mathbf{w} \qquad \text{where} \qquad \mathbf{M} = \mathbf{m}\cdot\mathbf{m}^H$$

The maximizing solution is $\mathbf{w} = c\cdot\mathbf{m}$ with $\lambda = \dfrac{\mathbf{m}^H\cdot\mathbf{m}}{\sigma^2}$ and maximized SNR

$$\gamma_{max} = \frac{\mathbf{m}^H\cdot\mathbf{m}}{\sigma^2}$$

(b) If the individual variances are different, then

$$\gamma = \frac{\mathbf{w}^H\cdot\mathbf{m}\cdot\mathbf{m}^H\cdot\mathbf{w}}{\mathbf{w}^T\cdot\mathbf{S}\cdot\mathbf{w}} \qquad \text{which is maximized with} \qquad \mathbf{w_{opt}} = \mathbf{S}^{-1}\cdot\mathbf{m}$$

using the same arguments as in the earlier solution. That is, the optimum weight vector has coeff

$$w_{opt_i} = \frac{m_i}{(\sigma_i)^2}$$

and we form the sum

$$d = \sum_{i=1}^{N} \frac{\overline{m_i}}{(\sigma_i)^2}\cdot x_i$$

Here's an interpretation: the conjugation of the mean value will "derotate" the mean value compo of x_i, making a real product; consequently those contributions to the sum add constructively, with internal cancellations; meanwhile, the statistics of the noise components are unaffected by the conjugation of the weight.

What is the resulting maximized SNR? Substitute this optimized choice for **w** back into the expression for γ and we get

$$\gamma = \frac{\mathbf{w_{opt}}^T\cdot\mathbf{m}\cdot\mathbf{m}^T\cdot\mathbf{w_{opt}}}{\mathbf{w_{opt}}^T\cdot\mathbf{S}\cdot\mathbf{w_{opt}}} = \frac{\left(\mathbf{m}^H\cdot\mathbf{S}^{-1}\cdot\mathbf{m}\right)^2}{\mathbf{m}^H\cdot\mathbf{S}^{-H}\cdot\mathbf{S}\cdot\mathbf{S}^{-1}\cdot\mathbf{m}} = \mathbf{m}^H\cdot\mathbf{S}^{-1}\cdot\mathbf{m}$$

since **S** is symmetric.