



answer questions



⊗ Thought for 74 seconds ∨

- First, I need to find the Wiener-Hopf solution for two different cost functions. The Wiener-Hopf solution typically refers to the optimal filter weights that minimize the mean square error, often in the context of linear estimation. Here, the cost functions include an expectation of the squared error plus a regularization term.

Let me recall the standard Wiener-Hopf equation. For a system where we have a desired signal  $y$  and an input matrix  $X$ , the optimal weights  $w$  without regularization satisfy  $E[X X^T] w = E[X y]$ . With regularization, it might change slightly.

Now, looking at the first cost function:

$$1. J(\bar{w}) = \mathbb{E} \|\bar{y} - X^T \bar{w}\|_2^2 + \lambda \|\bar{w}\|_2^2$$

I need to find the  $w$  that minimizes this. Since it's a quadratic form, I can take the derivative and set it to zero.

Let me denote  $w$  as the weight vector, and assume that  $X$  is a matrix where each column is an input vector, so  $X^T w$  is the output. But in the error, it's  $\bar{y} - X^T \bar{w}$ , so  $y$  is a vector, and  $X^T w$  is a vector. The expectation is over the random variables.

To minimize  $J(w)$ , I can expand the terms.

Let me write  $J(w)$  as:

$$J(w) = \mathbb{E} [(\bar{y} - X^T \bar{w})^T (\bar{y} - X^T \bar{w})] + \lambda \bar{w}^T \bar{w}$$

Since the L2 norm squared is the inner product.

$$\text{So, } J(w) = \mathbb{E} [\bar{y}^T \bar{y} - 2\bar{y}^T X^T \bar{w} + \bar{w}^T X X^T \bar{w}] + \lambda \bar{w}^T \bar{w}$$

Now, take the expectation inside:

$$J(w) = \mathbb{E}[\bar{y}^T \bar{y}] - 2\mathbb{E}[\bar{y}^T X^T] \bar{w} + \bar{w}^T \mathbb{E}[X X^T] \bar{w} + \lambda \bar{w}^T \bar{w}$$

I assume that the expectation is with respect to the joint distribution of  $y$  and  $X$ . So, let me define:

Let  $R_{xx} = \mathbb{E} [X X^T]$ , which is the autocorrelation matrix of  $X$ .

And let  $p = \mathbb{E} [X \overline{y}]$ , since  $\mathbb{E} [\overline{y}^T X^T] = p^T$ , but since we have  $\overline{y}^T X^T$ , it's a scalar, so  $\mathbb{E} [\overline{y}^T X^T] =$

$$\left( \mathbb{E} [ X \overline{y} ] \right)^T = p^T.$$

Actually, in the term, it's  $\overline{y}^T X^T \overline{w}$ , which is a scalar, so  $\mathbb{E} [ \overline{y}^T X^T \overline{w} ] = \left( \mathbb{E} [ X \overline{y} ] \right)^T \overline{w} = p^T \overline{w}$ .

Similarly,  $\mathbb{E} [ \overline{y}^T \overline{y} ]$  is a scalar, but it doesn't depend on  $w$ , so for minimization, it's constant.

$$\text{So, } J(w) = c - 2 p^T w + w^T R_{xx} w + \lambda w^T w$$

$$\text{where } c = \mathbb{E} [ \overline{y}^T \overline{y} ]$$

Now, to find the minimum, take the gradient with respect to  $w$  and set to zero.

$$\nabla_w J(w) = -2 p + 2 R_{xx} w + 2 \lambda w = 0$$

$$\text{So, } -2 p + 2 ( R_{xx} w + \lambda w ) = 0$$

$$\text{Thus, } R_{xx} w + \lambda w = p$$

$$\text{So, } ( R_{xx} + \lambda I ) w = p$$

$$\text{Therefore, the optimal } w \text{ is } w^* = ( R_{xx} + \lambda I )^{-1} p$$

This is the Wiener-Hopf solution for the first cost function.

Now, for the second cost function:

$$2. J(\overline{w}) = \mathbb{E} \| \overline{y} - X^T \overline{w} \|_2^2 + \lambda \| \overline{w} - \overline{w}^s \|_2^2$$

Here, we have a regularization term that pulls  $w$  towards a prior weight vector  $\overline{w}^s$ .

Similarly, I can expand this.

$$J(w) = \mathbb{E} \left[ (\overline{y} - X^T \overline{w})^T (\overline{y} - X^T \overline{w}) \right] + \lambda (\overline{w} - \overline{w}^s)^T (\overline{w} - \overline{w}^s)$$

$$= \mathbb{E} [ \overline{y}^T \overline{y} - 2 \overline{y}^T X^T \overline{w} + \overline{w}^T X X^T \overline{w} ] + \lambda [ \overline{w}^T \overline{w} - 2 \overline{w}^T \overline{w}^s + (\overline{w}^s)^T \overline{w}^s ]$$

Now, take expectation:

$$J(w) = \mathbb{E} [ \overline{y}^T \overline{y} ] - 2 \mathbb{E} [ \overline{y}^T X^T \overline{w} ] + \overline{w}^T \mathbb{E} [ X X^T ] \overline{w} + \lambda \overline{w}^T \overline{w} - 2 \lambda \overline{w}^T \overline{w}^s + \lambda (\overline{w}^s)^T \overline{w}^s$$

$$\text{Again, let } c = \mathbb{E} [ \overline{y}^T \overline{y} ], R_{xx} = \mathbb{E} [ X X^T ], p = \mathbb{E} [ X \overline{y} ]$$

$$\text{So, } J(w) = c - 2 p^T w + w^T R_{xx} w + \lambda w^T w - 2 \lambda w^T \overline{w}^s + \lambda (\overline{w}^s)^T \overline{w}^s$$

Now, take gradient with respect to  $w$ :

$$\nabla_w J(w) = -2 p + 2 R_{xx} w + 2 \lambda w - 2 \lambda \overline{w}^s = 0$$

$$\mathbb{E}[\overline{y} \overline{w}^s]$$

Again, let  $c = \mathbb{E}[\overline{y}^T \overline{y}]$ ,  $R_{xx} = \mathbb{E}[X X^T]$ ,  $p = \mathbb{E}[X \overline{y}]$

So,  $J(w) = c - 2 p^T w + w^T R_{xx} w + \lambda w^T w - 2 \lambda w^T w^s + \lambda (w^s)^T w^s$

Now, take gradient with respect to  $w$ :

$$\nabla_w J(w) = -2 p + 2 R_{xx} w + 2 \lambda w - 2 \lambda w^s = 0$$

So, to summarize:

For cost function 1:  $w^* = (R_{xx} + \lambda I)^{-1} p$

For cost function 2:  $w^* = (R_{xx} + \lambda I)^{-1} (p + \lambda w^s)$

Where  $R_{xx} = \mathbb{E}[X X^T]$  and  $p = \mathbb{E}[X \overline{y}]$

I think that's the Wiener-Hopf solution.

Now, I need to present this in the answer.

The assignment says "find the Wiener-Hopf solution", so I should state these equations.

Also, in the cost function, it's written with overlines, so I'll keep the notation consistent.

In the first cost function, it's  $\overline{y}$  and  $\overline{w}$ , but in the second, it's  $\overline{w}$  and  $\overline{w}^s$ , so I'll use  $w$  for  $\overline{w}$  for simplicity.

So, for the answer, I'll write:

The Wiener-Hopf solution for the given cost functions is derived by minimizing the cost function with respect to the weight vector  $\overline{w}$ . The solutions are as follows:

**1. For the cost function  $J(\overline{w}) = \mathbb{E} \|\overline{y} - X^T \overline{w}\|_2^2 + \lambda \|\overline{w}\|_2^2$**

The optimal weight vector  $\overline{w}^*$  is given by:

$$\overline{w}^* = (R_{xx} + \lambda I)^{-1} p$$

where:

- $R_{xx} = \mathbb{E}[X X^T]$  is the autocorrelation matrix of  $X$ ,
- $p = \mathbb{E}[X \overline{y}]$  is the cross-correlation vector between  $X$  and  $\overline{y}$ ,
- $I$  is the identity matrix.

**2. For the cost function  $J(\overline{w}) = \mathbb{E} \|\overline{y} - X^T \overline{w}\|_2^2 + \lambda \|\overline{w} - \overline{w}^s\|_2^2$**

The optimal weight vector  $\overline{w}^*$  is given by:

$$\bar{w}^* = (R_{xx} + \lambda I)^{-1}(p + \lambda \bar{w}^s)$$

where:


- $R_{xx} = \mathbb{E}[XX^T]$  is the autocorrelation matrix of  $X$ ,
- $p = \mathbb{E}[X\bar{y}]$  is the cross-correlation vector between  $X$  and  $\bar{y}$ ,
- $\bar{w}^s$  is the prior weight vector,
- $I$  is the identity matrix.

These solutions are obtained by taking the gradient of the cost function with respect to  $\bar{w}$ , setting it to zero, and solving for  $\bar{w}$ . The regularization parameter  $\lambda$  controls the trade-off between the error term and the regularization term.



Message DeepSeek

 DeepThink

 Search



AI-generated, for reference only