EECS 844 – Fall 2016

Exam 1 Cover page*

Each student is expected to complete the exam individually using only course notes, the book, and technical literature, and without aid from outside sources.

Aside from the most general conversation of the exam material, I assert that I have neither provided help nor accepted help from another student in completing this exam. As such, the work herein is mine and mine alone.

Signature ______________________ Date ________

MY SOLUTIONS

Name (printed) ______________________ Student ID # ________

*** Attach as cover page to completed exam.**

1. For the cost functions below, determine the derivative with respect to $\mathbf{w}^*$. (*Note: the chain rule still holds*)

$$\text{a)}\quad J(\mathbf{w}) = \exp\left(-\mathbf{w}^H \mathbf{R}\mathbf{w}\right)$$

$$\text{b)}\quad J(\mathbf{w}) = \frac{1}{\|\mathbf{w}\|}$$

$$\text{c)}\quad J(\mathbf{w}) = \text{trace}\left\{\mathbf{w}\mathbf{w}^H\right\}$$

$$\text{d)}\quad J(\mathbf{w}) = \frac{\left|\mathbf{w}^H \mathbf{a}\right|^2}{\mathbf{w}^H \mathbf{R}\mathbf{w}} \text{ for } \mathbf{R} = \mathbf{R}^H$$

Solutions:

Comment 1: See Sect. 4.1 of 2012 edition of *The Matrix Cookbook* (available free online and I encourage you all to get a copy). The complex chain rule is the standard chain rule if involving analytic functions (which means the function is complex differentiable). In short, we don't need to consider the complicated form of the chain rule for these problems.

Comment 2: Use parentheses in places where two adjacent vectors are multiplied but have incompatible dimensions. For example, $(\mathbf{b}^H\mathbf{w})\mathbf{b}$ is correct but in $\mathbf{b}^H\mathbf{w}\mathbf{b}$ the multiplication of the last two vectors cannot be performed due to dimensionality. Of course, rearranging as $\mathbf{b}\mathbf{b}^H\mathbf{w}$ is a feasible multiplication without parentheses. Use this formalism to be technically correct and also to avoid possible mistakes.

Comment 3: When expressing math via a word processor (like you would in a paper as opposed to hand written), make scalars italic & lowercase, make vectors bold & lowercase, and make matrices bold & uppercase. Vectors and matrices are not italicized according to the general signal processing convention.

a)

$$\begin{aligned}
\frac{dJ(\mathbf{w})}{d\mathbf{w}^*} &= \frac{d}{d\mathbf{w}^*}\left[\frac{1}{a^{\left(0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}}\right] \\
&= a^{\left(-0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)} \\
&= -0.5\ln(a)\ a^{\left(-0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}\mathbf{R}\mathbf{w}
\end{aligned}$$

b)

$$\begin{aligned}\frac{dJ(\mathbf{w})}{d\mathbf{w}^*} &= \frac{d}{d\mathbf{w}^*}\left[\|\mathbf{w}\|^3\right]\\ &= \frac{d}{d\mathbf{w}^*}\left[\left(\mathbf{w}^H\mathbf{w}\right)^{3/2}\right]\\ &= 1.5\mathbf{w}\left(\mathbf{w}^H\mathbf{w}\right)^{1/2}\\ &= \left(1.5\|\mathbf{w}\|\right)\mathbf{w}\end{aligned}$$

c) First note that

$$\mathrm{Re}\left\{\mathbf{w}^H\mathbf{b}\right\} = 0.5\left(\mathbf{w}^H\mathbf{b} + \mathbf{b}^H\mathbf{w}\right)$$

Thus

$$\begin{aligned}\frac{d}{dw_m^*}\left[\mathrm{Re}\left\{\mathbf{w}^H\mathbf{b}\right\}\right] &= \frac{d}{dw_m^*}\left[0.5\left(\mathbf{w}^H\mathbf{b} + \mathbf{b}^H\mathbf{w}\right)\right]\\ &= 0.5\ \mathbf{b}\end{aligned}$$

d)

$$\frac{dJ(\mathbf{w})}{d\mathbf{w}^*} = \frac{d}{d\mathbf{w}^*}\left[\frac{\left|\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\right|^2}{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}\right]\frac{1}{\left(\mathbf{a}^H\mathbf{a}\right)} = \frac{d}{d\mathbf{w}^*}\left[\frac{\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w}}{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}\right]\frac{1}{\left(\mathbf{a}^H\mathbf{a}\right)}$$

$$= \frac{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)\frac{d}{d\mathbf{w}^*}[\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w}] \;-\; (\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w})\frac{d}{d\mathbf{w}^*}[\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)]}{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)^2\left(\mathbf{a}^H\mathbf{a}\right)}$$

$$= \frac{(\mathbf{w}^H\mathbf{R}\mathbf{w})\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w} \;-\; (\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w})\mathbf{R}\mathbf{w}}{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)^2\left(\mathbf{a}^H\mathbf{a}\right)}$$

2. For the time-series data in P2.mat, estimate the "temporal" correlation matrix **R** by forming a matrix **X** of delay-shifted snapshots (each of length M) so that $\mathbf{R} = (1/N)\,\mathbf{X}\mathbf{X}^H$, where N is the number of columns in **X** (see Appendix A). Using the **R** estimate for $M = 8$,

a) plot the eigenvalues of **R** (in dB)
b) determine the condition number
c) compute $\mathbf{R}^{-1}$ and plot its eigenvalues (in dB)
d) how are the two sets of eigenvalues related?

<u>Solution:</u>
The eigenvalues of **R** and $\mathbf{R}^{-1}$ are plotted in Fig. 2.1. As expected the two sets of eigenvalues are the inverse of one another. The condition number for this correlation matrix is found to be roughly 2.7×10^3. This large value is expected since the eigenvalues have a maximum value of about +13 dB and a minimum value of about –21 dB. This correlation matrix would likely have noise enhancement problems as demonstrated by the small eigenvalues becoming large when inverted.

Comment 1: Note that the 'eig' function in Matlab often provides the eigenvalues in descending order (with associated eigenvectors ordered accordingly). As shown below, each eigenvalue of **R** and $\mathbf{R}^{-1}$ correspond to the same eigenvector.

Comment 2: Note that correlation and covariance are associated with power (not amplitude) so you would plot eigenvalues using 10 log and not 20 log.

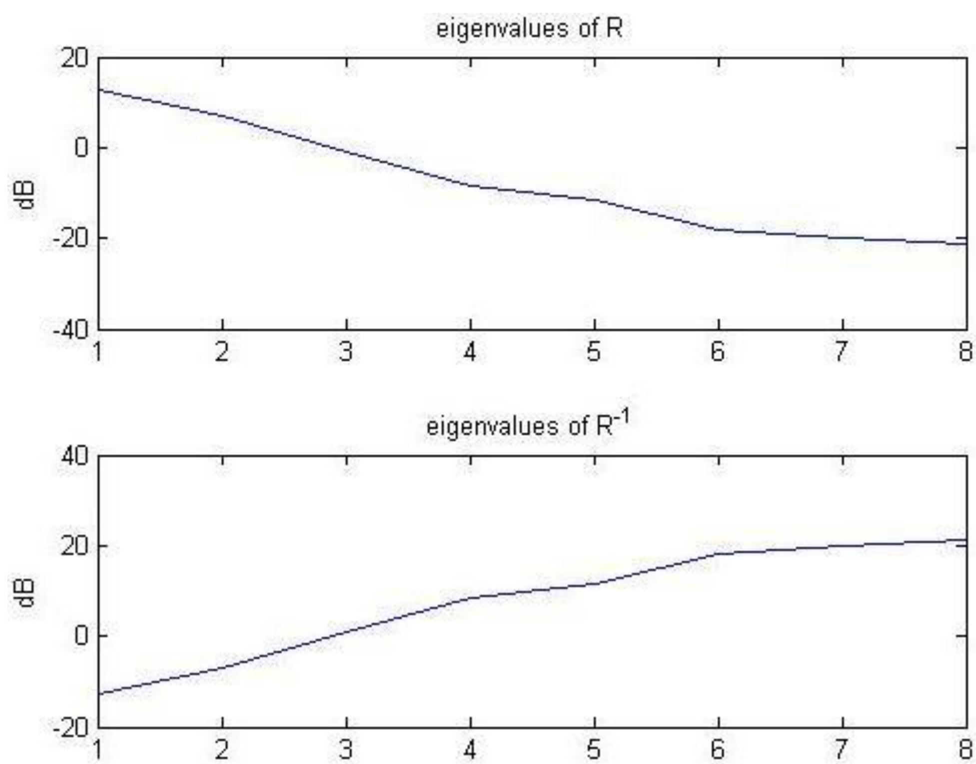

Figure 2.1. Eigenvalues of **R** and $\mathbf{R}^{-1}$

Matlab Code for Problem 2

```
clear all;
load P2
N = length(x);

M = 8;

for jm = 1:M
    X(M-jm+1,:) = x(jm:jm+N-M).';
end;

R = (1./(N-M+1)).*X*X';

[V,D] = eig(R);
cond(R)

figure(21)
subplot(2,1,1)
plot(1:M,10*log10(diag(D)));
ylabel('dB')
title('eigenvalues of R')
subplot(2,1,2)
plot(1:M,10*log10(diag(D^(-1))));
ylabel('dB')
title('eigenvalues of R^{-1}')
```

3. Repeat problem 2 using the "diagonally loaded" correlation matrix estimate defined as $\mathbf{R} = (1/N)\, \mathbf{X}\mathbf{X}^H + \sigma^2\mathbf{I}$, where $\sigma^2\mathbf{I}$ is the correlation matrix of white noise. Here set $\sigma^2 = 1$. What do you observe in comparison to the Problem 2 results?

<u>Solution:</u>
The eigenvalues of the diagonally loaded versions of $\mathbf{R}$ and $\mathbf{R}^{-1}$ are plotted in Fig. 3.1 (in blue) along with the previous result from prob. 2 (in green). The two sets of eigenvalues are still the inverse of one another. The condition number for the diagonally loaded correlation matrix is found to be roughly 20.4, noticeably less than for Prob. 2. This reduction in the condition number is realized because the small eigenvalues now do not go below 0 dB (blue curve). As a result, when inverted the noise enhancement effect would no longer occur.

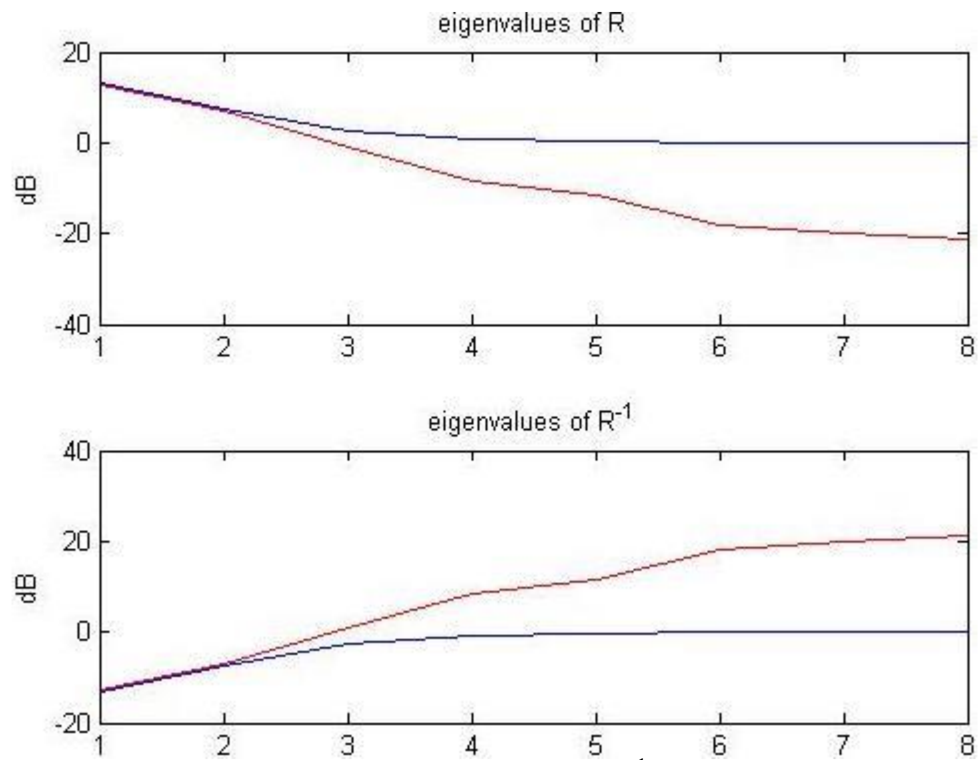

Figure 3.1. Eigenvalues of diagonally loaded $\mathbf{R}$ and $\mathbf{R}^{-1}$. Diagonally loaded version is in red, original is in blue.

It can be shown that the eigenvectors for the diagonally loaded correlation matrix are identical to the eigenvectors of the original correlation matrix from Prob. 2.

<u>Matlab Code for Problem 3</u>

```
clear all;
load P2
N = length(x);

M = 8;

for jm = 1:M
    X(M-jm+1,:) = x(jm:jm+N-M).';
end;

R = (1./(N-M+1)).*X*X';
[V,D] = eig(R);

sig = 1;
Rd = (1./(N-M+1)).*X*X' + sig.*eye(M);

[Vd,Dd] = eig(Rd);
cond(Rd)

V3V2 = Vd'*V;

figure(31)
subplot(2,1,1)
plot(1:M,10*log10(diag(Dd)),1:M,10*log10(diag(D)));
ylabel('dB')
title('eigenvalues of R')
subplot(2,1,2)
plot(1:M,10*log10(diag(Dd^(-1))),1:M,10*log10(diag(D^(-1))));
ylabel('dB')
title('eigenvalues of R^{-1}')
```

4. For the correlation matrix estimate **R** from Problem 2 with associated eigenvector matrix **V**, transform the data matrix as $\mathbf{X}_{new} = \mathbf{V}^H\mathbf{X}$. What do you observe about the correlation matrix of this new transformed data? (show plots as appropriate)

Solution:

Using the transformed data $\mathbf{X}_{new}$ the resulting new covariance matrix $\mathbf{R}_{new}$ is a diagonal matrix whose diagonal values equal the eigenvalues of the original correlation matrix from Prob. 2. In other word, $\mathbf{R}_{new} = \mathbf{\Lambda}$. Figure 4.1 shows the eigenvalues of the original **R** and the diagonal values of $\mathbf{R}_{new}$. Using the 'imagesc' command to plot the magnitude (in dB) of all the elements of the $\mathbf{R}_{new}$ matrix, Fig. 4-2 shows that all the off-diagonal elements are essentially zero.

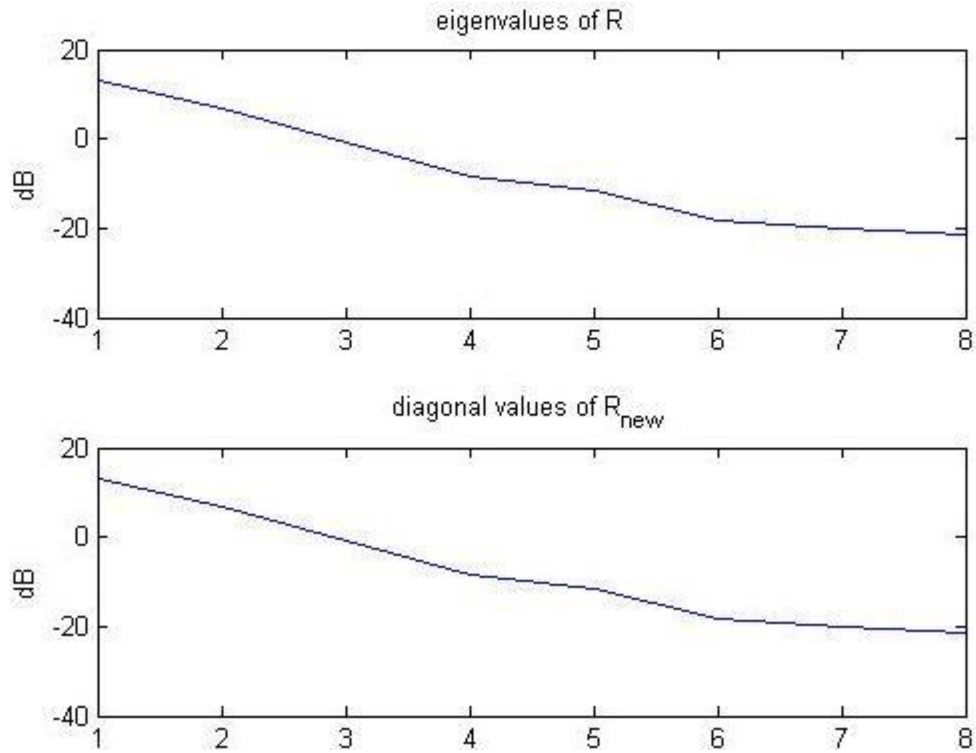

Figure 4.1. Eigenvalues of **R** and diagonal values of $\mathbf{R}_{new}$

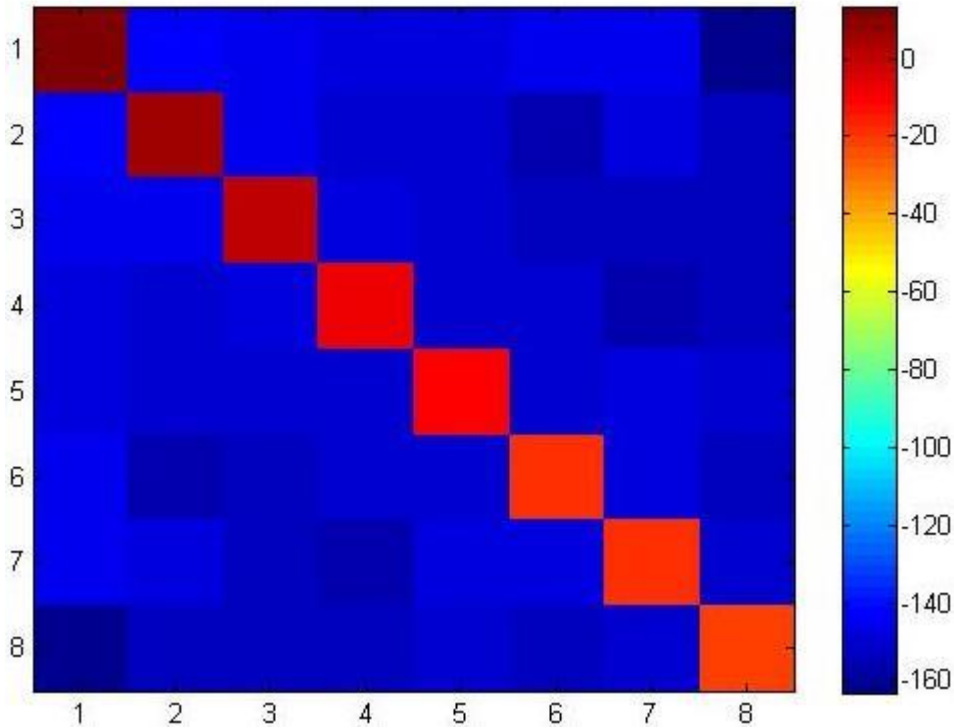

Figure 4.2. dB plot of the $\mathbf{R}_{\text{new}}$ matrix magnitudes

The reason for this effect can be observed when considering the math. Recall that the correlation matrix is

$$\mathbf{R} = \frac{1}{N}\mathbf{X}\mathbf{X}^H = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H .$$

So when the data is transformed as $\mathbf{X}_{\text{new}} = \mathbf{V}^H\mathbf{X}$, the new correlation matrix is

$$\begin{aligned}
\mathbf{R}_{\text{new}} &= \frac{1}{N}\mathbf{X}_{\text{new}}\mathbf{X}_{\text{new}}^H \\
&= \frac{1}{N}(\mathbf{V}^H\mathbf{X})(\mathbf{V}^H\mathbf{X})^H \\
&= \frac{1}{N}\mathbf{V}^H\mathbf{X}\mathbf{X}^H\mathbf{V} \\
&= \frac{N}{N}\mathbf{V}^H\mathbf{R}\mathbf{V} \\
&= \mathbf{V}^H(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H)\mathbf{V} \\
&= \mathbf{\Lambda}
\end{aligned}$$

which is what is observed in Figs. 4.1 and 4.2.

Matlab Code for Problem 4

```
clear all;
load P2
N = length(x);

M = 8;

for jm = 1:M
    X(M-jm+1,:) = x(jm:jm+N-M).';
end;

R = (1./(N-M+1)).*X*X';
[V,D] = eig(R);

Xt = V'*X;

Rt = (1./(N-M+1)).*Xt*Xt';

figure(41)
subplot(2,1,1)
plot(1:M,10*log10(diag(D)));
ylabel('dB')
title('eigenvalues of R')
subplot(2,1,2)
plot(1:M,10*log10(diag(Rt)));
ylabel('dB')
title('diagonal values of Rt')

figure(42)
imagesc(10*log10(abs(Rt)))
colorbar
```

5. The dataset P5.mat contains time samples collected from an $M = 12$ element antenna array. Using this data, estimate the "spatial" correlation matrix $\mathbf{R}$ and subsequently

a) plot the eigenvalues (in dB)
b) determine the condition number
c) compute $\mathbf{R}^{-1}$ and plot its eigenvalues (in dB)
d) discuss how are the two sets of eigenvalues related.

Solution:

The spatial covariance matrix corresponding to a 12 element antenna array would have dimensionality 12×12 due to the physical spatial dimensionality. The provided matrix $\mathbf{X}$ is 12×100, with each column representing the samples captured by the array at a given time instant. In other words, $\mathbf{X}$ is oriented as "number of antenna elements" × "number of time samples". Therefore, computing $\mathbf{R} = (1/N)\,\mathbf{X}\mathbf{X}^H$ produces the spatial correlation matrix. The eigenvalues of $\mathbf{R}$ and $\mathbf{R}^{-1}$ are plotted in Fig. 5.1. As expected the two are again the inverse of one another. The condition number is found to be roughly 4.8×10^3. The relatively flat region represents the "noise subspace" while the rest corresponds to the "(signal + noise) subspace". We can infer (perhaps correctly as will be discussed later) that there are 6 signals present in the data as that is the number of eigenvalues that stand out relative to the noise.

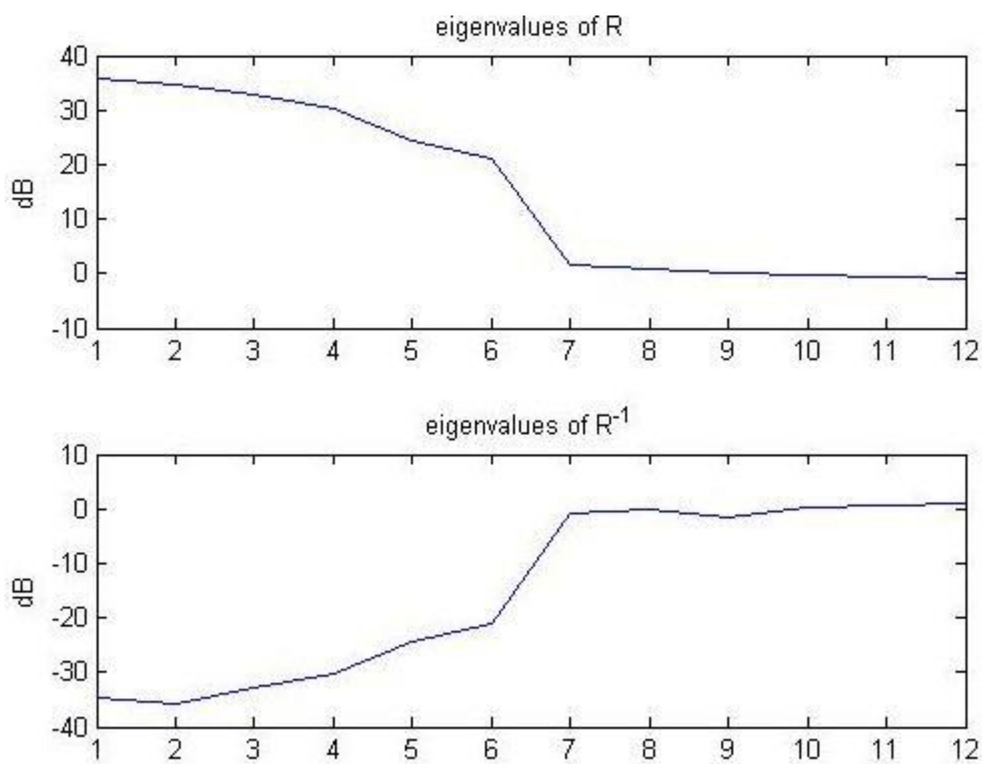

Figure 5.1. Eigenvalues of $\mathbf{R}$ and $\mathbf{R}^{-1}$

Matlab Code for Problem 5

```
clear all
load P5

[M N] = size(X);

R = X*X'./N;
[V,D] = eig(R);
cond(R)

figure(1)
subplot(2,1,1)
plot(1:M,10*log10(flipud(diag(D))));
axis([1 M -10 40])
ylabel('dB')
title('eigenvalues of R')
subplot(2,1,2)
plot(1:M,10*log10(flipud(diag(D^(-1)))));
axis([1 M -40 10])
ylabel('dB')
title('eigenvalues of R^{-1}')
```

6. The dataset P6.mat contains two length M = 12 antenna array filters (i.e. beamformers) for a uniform linear array whose elements are separated by a half-wavelength. The filter 'w_non_adap' is a non-adaptive filter while the filter 'w_adap' is an adaptive filter. Using Appendices B and C, plot the beampatterns of these two filters in terms of electrical angle and spatial angle (plot in dB). Discuss what you observe.

<u>Solution:</u>
The beampatterns in terms of electrical angle and spatial angle are plotted in Figs. 6-1 and 6-2, respectively, where the blue trace is the non-adaptive filter response and the red trace is the adaptive filter response. The adaptive response appears to be poorer since it has slightly higher sidelobes. However, the data used to form the adaptive filter contains interference sources at electrical angles –63.3°, +42.1°, and +131.1°, for which Fig. 6-1 show nulls in the adaptive beampattern corresponding to these angles. The high sidelobes in the adaptive response do not degrade performance since no signals are presently arriving from those directions (or else they would have been present in the data). Of course, there is an assumption here that the data is sufficiently stationary that the interference does not change locations and new sources do not appear.

The nonlinear relationship between electrical angle and spatial angle (due to the sin(•) function) can also be observed. The electrical angle sidelobes are uniformly spaced while the spatial sidelobes are nonlinear in nature (narrower near boresight and wider near endfire).

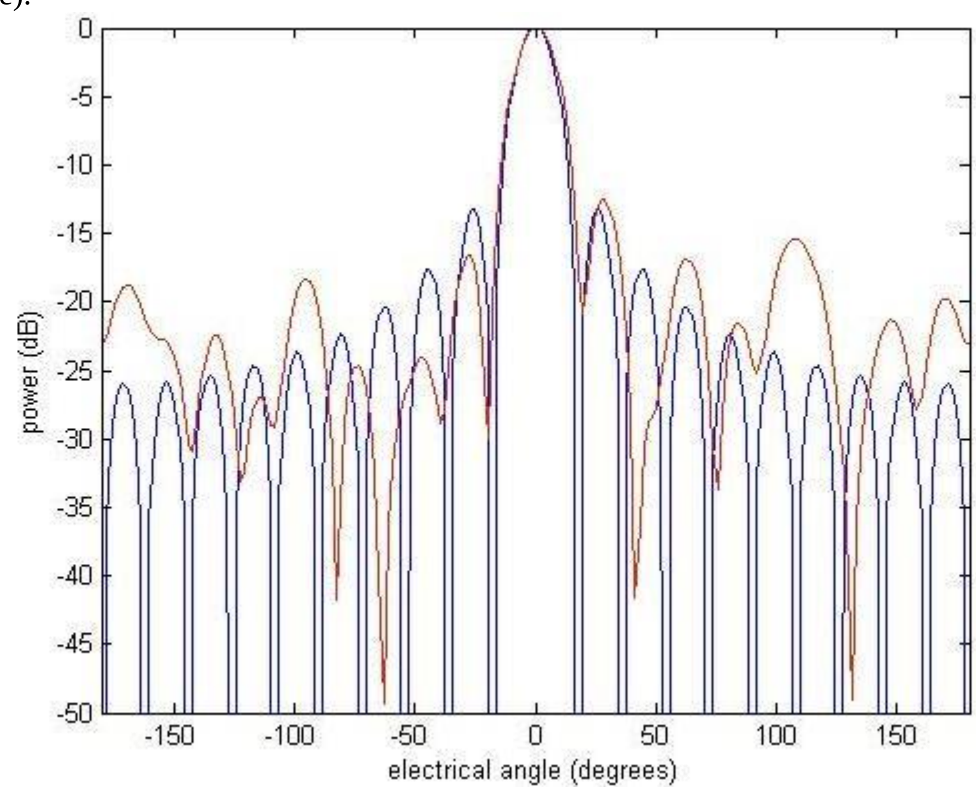

Figure 6.1. Beampatterns in terms of electrical angle ($-\pi \leq \theta \leq \pi$). Non-adaptive beamformer is blue, adaptive beamformer is red.

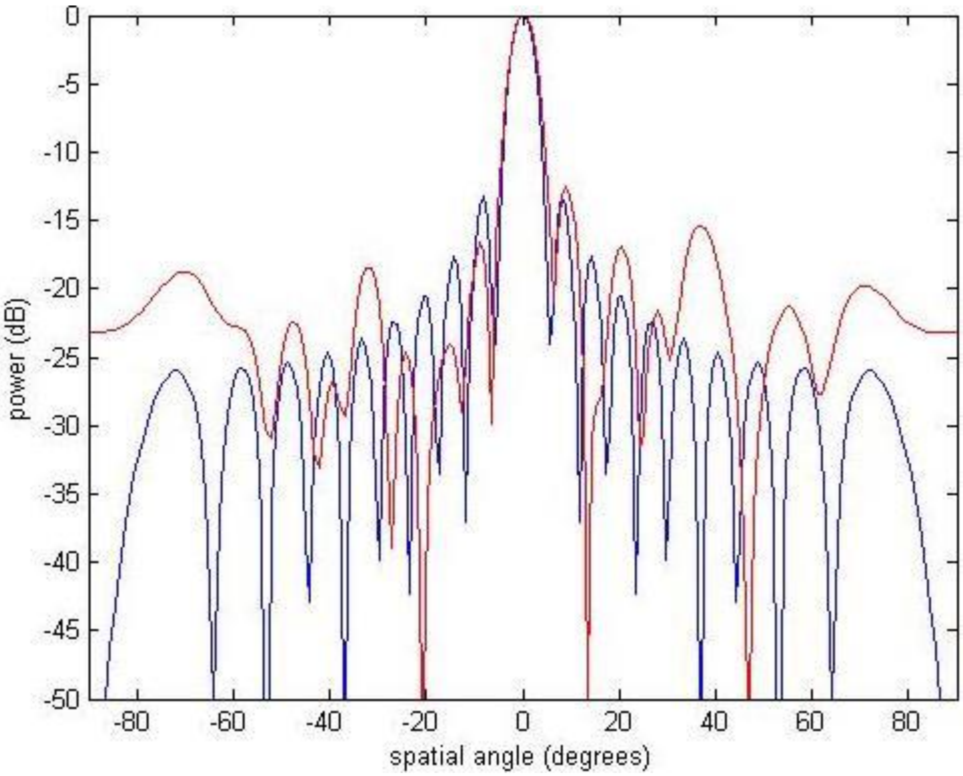

Figure 6.2. Beampatterns in terms of spatial angle ($-\pi/2 \leq \phi \leq \pi/2$). Non-adaptive beamformer is blue, adaptive beamformer is red.

<u>Matlab Code for Problem 6</u>

```
clear all
load P6

M = length(w_adap);

angs_spat = linspace(-90,90,M*10+1);
angs_elec = linspace(-180,180,M*10+1);

for theta_i = 1:length(angs_spat)
    s_spat = (exp(-j*pi*sin((angs_spat(theta_i)*pi/180)).*[0:M-1]).');
    s_elec = (exp(-j*(angs_elec(theta_i)*pi/180).*[0:M-1]).');
    S_spat(1:M,theta_i) = s_spat;
    S_elec(1:M,theta_i) = s_elec;
end;

pattern_nonadap_spat = abs(S_spat'*w_non_adap);
pattern_nonadap_elec = abs(S_elec'*w_non_adap);
pattern_adap_spat = abs(S_spat'*w_adap);
pattern_adap_elec = abs(S_elec'*w_adap);

figure(1)
plot(angs_spat,20*log10(pattern_nonadap_spat),angs_spat,20*log10(patter
n_adap_spat))
axis([-90 90 -50 0])
xlabel('spatial angle (degrees)')
ylabel('power (dB)')

figure(2)
plot(angs_elec,20*log10(pattern_nonadap_elec),angs_elec,20*log10(patter
n_adap_elec))
axis([-180 180 -50 0])
xlabel('electrical angle (degrees)')
ylabel('power (dB)')
```

7. For each of the cost functions in Problem 1, apply the linear constraint $\mathbf{w}^H\mathbf{s}=1$ and solve for the complex Lagrange multiplier. You can assume all matrices are PDH. (*In each case, pre-multiply by* $\mathbf{s}^H$ *when solving*)

<u>Solutions:</u>

For all four cases, define $c(\mathbf{w})=\mathbf{w}^H\mathbf{s}-1$.

Reminder: You <u>cannot</u> divide by a vector or matrix. To solve for λ (after taking the derivative) pre-multiply by $\mathbf{s}^H$ to obtain an inner product that results in scalar terms that can then be manipulated (i.e. divided by).

a) Supplement the cost function with the constraint as

$$h(\mathbf{w})=\frac{1}{a^{\left(0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}}+\lambda^*\left(\mathbf{w}^H\mathbf{s}-1\right)$$

so

$$\begin{aligned}\frac{dh(\mathbf{w})}{d\mathbf{w}^*}&=\frac{d}{d\mathbf{w}^*}\left[a^{\left(-0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}+\lambda^*\left(\mathbf{w}^H\mathbf{s}-1\right)\right]\\&=-0.5\ \ln(a)\ a^{\left(-0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}\mathbf{R}\mathbf{w}+\lambda^*\mathbf{s}=\mathbf{0}_{M\times 1}\end{aligned}.$$

Rearrange this equation, pre-multiply by $\mathbf{s}^H$, and then solve for the Lagrange multiplier:

$$\lambda=0.5\ \ln(a)\ a^{\left(-0.5\mathbf{w}^H\mathbf{R}\mathbf{w}\right)}\left(\frac{\mathbf{w}^H\mathbf{R}\mathbf{s}}{\mathbf{s}^H\mathbf{s}}\right)$$

b) Supplement the cost function with the constraint as

$$h(\mathbf{w}) = \|\mathbf{w}\|^3 + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)$$

so

$$\begin{aligned}\frac{dh(\mathbf{w})}{d\mathbf{w}^*} &= \frac{d}{d\mathbf{w}^*}\left[\left(\mathbf{w}^H \mathbf{w}\right)^{3/2} + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)\right] \\ &= \left(1.5\left(\mathbf{w}^H \mathbf{w}\right)^{1/2}\right)\mathbf{w} + \lambda^* \mathbf{s} = \mathbf{0}_{M\times 1}\end{aligned}$$

Rearrange this equation, pre-multiply by $\mathbf{s}^H$, and then solve for the Lagrange multiplier:

$$\lambda = \frac{-1.5\left(\mathbf{w}^H \mathbf{w}\right)^{1/2} \mathbf{w}^H \mathbf{s}}{\left(\mathbf{s}^H \mathbf{s}\right)} = \frac{-1.5\left(\mathbf{w}^H \mathbf{w}\right)^{1/2} (1)}{\left(\mathbf{s}^H \mathbf{s}\right)}$$

c) Supplement the cost function with the constraint as

$$h(\mathbf{w}) = \mathrm{Re}\left\{\mathbf{w}^H \mathbf{b}\right\} + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)$$

so

$$\begin{aligned}\frac{dh(\mathbf{w})}{d\mathbf{w}^*} &= \frac{d}{d\mathbf{w}^*}\left[\mathrm{Re}\left\{\mathbf{w}^H \mathbf{b}\right\} + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)\right] \\ &= 0.5\mathbf{b} + \lambda^* \mathbf{s} = \mathbf{0}_{M\times 1}\end{aligned}$$

Rearrange this equation, pre-multiply by $\mathbf{s}^H$, and then solve for the Lagrange multiplier:

$$\lambda = \frac{-0.5\,\mathbf{b}^H \mathbf{s}}{\left(\mathbf{s}^H \mathbf{s}\right)}$$

d) Supplement the cost function with the constraint as

$$h(\mathbf{w}) = \frac{\left|\mathbf{w}^H \mathbf{R}^{-1}\mathbf{a}\right|^2}{\left(\mathbf{w}^H \mathbf{R}\mathbf{w}\right)\left(\mathbf{a}^H \mathbf{a}\right)} + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)$$

so

$$\frac{dh(\mathbf{w})}{d\mathbf{w}^*} = \frac{d}{d\mathbf{w}^*}\left[\frac{\left|\mathbf{w}^H \mathbf{R}^{-1}\mathbf{a}\right|^2}{\left(\mathbf{w}^H \mathbf{R}\mathbf{w}\right)\left(\mathbf{a}^H \mathbf{a}\right)} + \lambda^* \left(\mathbf{w}^H \mathbf{s} - 1\right)\right]$$

$$= \frac{(\mathbf{w}^H\mathbf{R}\mathbf{w})\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w} \; - \; (\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w})\mathbf{R}\mathbf{w}}{\left(\mathbf{w}^H\mathbf{R}\mathbf{w}\right)^2\left(\mathbf{a}^H\mathbf{a}\right)} + \lambda^*\mathbf{s} = \mathbf{0}_{M\times 1}$$

Rearrange this equation, pre-multiply by $\mathbf{s}^H$, and then solve for the Lagrange multiplier:

$$\lambda = \frac{(\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a}\mathbf{a}^H\mathbf{R}^{-1}\mathbf{w})(\mathbf{w}^H\mathbf{R}\mathbf{s}) \; - \; (\mathbf{w}^H\mathbf{R}\mathbf{w})(\mathbf{a}^H\mathbf{R}^{-1}\mathbf{s}\mathbf{w}^H\mathbf{R}^{-1}\mathbf{a})}{(\mathbf{w}^H\mathbf{R}\mathbf{w})^2(\mathbf{a}^H\mathbf{a})(\mathbf{s}^H\mathbf{s})}$$