EECS120 - Fall 2003
Homework No. 4
Due: In the 120 Box On Sep 30 **You are strongly urged to turn in a photocopy of your work so that you can check against the solutions before the exam.**
Collaboration permitted and solutions to be written up by groups up to 3.
Be clear and precise in your answers
Questions can be asked in ucb.class.ee120 or in office hours

Problem 4.1 *Book Problems from Lee and Varaiya, chapter 10.*
Problems: 11, 13, 14, 15

Problem 4.2 *Projections and Least Squares.*
This problem is designed to help you review your linear algebra by exploring the special properties of the projection operation and see how it relates to the Discrete Fourier Transform and Fourier Series[1]

a. *One dimensional projections. Given an n-dimensional column vector $\vec{v}$, derive a matrix L that projects column vectors onto $\vec{v}$. (i.e. If $\vec{y} = L\vec{x}$, then $\vec{y}$ is the projection of $\vec{x}$ in the direction of $\vec{v}$.)*

b. *You perform an experiment that yields a column vector of n measurements denoted by $\vec{d}$. You believe this data to come from a physical process that should give measurements of the form $\alpha\vec{a}$ where α is some unknown scalar gain parameter. In order to estimate α, you decide to minimize the standard Euclidean norm of the error vector $\vec{e} = \vec{d} - \alpha\vec{a}$ where α is the parameter you get to set.*

   Show that the optimal α gives rise to $\alpha\vec{a}$ being the projection of $\vec{d}$ onto the direction given by $\vec{a}$.

   (HINT: Take the derivative of the squared norm of the error vector and set it to zero. Then show that this is the global minimum.)

c. *Show that the error vector $\vec{e}$ is orthogonal to $\vec{a}$.*

d. *Repeat part (b) above, but now your model for the measurements has two unknowns (α, β) and is of the form $\alpha\vec{a} + \beta\vec{b}$ where $\vec{a}$ and $\vec{b}$ are linearly independent. What is the choice of α and β that minimizes the norm of the error vector?*

e. *Show that the error vector from part (d) is orthogonal to both $\vec{a}$ and $\vec{b}$ and hence to the entire subspace spanned by the two of them.*

---

[1]You are urged to explore things in low-dimensional spaces like two or three dimensions before solving the general n-dimensional cases asked for below. If you have questions on how to prove things or do these problems, see us in office hours.

f. *Repeat part (d) for the case of a model with m linearly independent vectors $\vec{a}_i$ and m corresponding unknowns α_i.*

(HINT: Put the m column vectors $\vec{a}_i$ into a matrix A and then collect all the α_i unknown parameters into a column vector $\vec{x}$. Now the linear combination is given by the matrix multiplication $A\vec{x}$ and the error is $\vec{d} - A\vec{x}$. Calculate the norm squared and differentiate with respect to $\vec{x}$ to help find the global optimum choice of $\vec{x}$)

g. *Now, suppose in part (f) above that the vectors $\vec{a}_i$ are all orthogonal to each other. Simplify your answer to part (f) for this case. What is special about orthogonality?*

h. *BONUS: Show that the error vector in part (f) is orthogonal to the entire subspace spanned by the $\vec{a}_i$.*

(HINT: You can do this very easily by using part (g) and feeding it a set of orthogonal vectors generated from the original $\vec{a}_i$'s by using Gramm-Schmidt on them. (You don't actually have to do it, just realize that you can and that it would give you m orthogonal vectors spanning the same subspace.) Then, throw in another $m + 1$-th orthogonal vector by adding $\vec{d}$ to the end of the list for Gramm-Schmidt orthogonalization. Now apply (g) twice: once with the m orthogonal vectors, and once again with the $m + 1$ of them and interpret what happens.)

i. *Interpret (g) and extend it to the case of trying to model a finite duration continuous time complex signal $s(t)$ (defined for $t \in [0, T]$ you can also interpret this as a periodic signal with period T and defined for all real t) as a weighted sum of the first $2m+1$ of the T-periodic complex exponentials $e^{j\frac{2\pi i}{T}t}$, for $i = -m, -m+1, \ldots, -1, 0, 1, \ldots, +m$. How would you pick the α_i? Does your choice minimize the energy in the error?*

(HINT: Think about the results of (g) in terms of inner products and norms as we did in class. This is related to the discrete Fourier series.)

j. *BONUS: Suppose that you wished to represent a finite duration discrete time signal $s(t)$ with $t \in \{0, 1, \ldots, T-1\}$ but instead of using the first $2m + 1$ of the T-periodic complex exponentials ($2m + 1 < T$), you decided to use the $2m + 1$ of them that gave the smallest energy in the error. Give an algorithm and justification for why it works. (Variations of this are actually used in doing lossy image and audio compression by combining this idea with another norm that accounts for perceptual effects.)*

(HINT: Think about using the DFT and relating it to the approach in part (f). To help you understand what is going on, experiment with various $s(t)$s in MATLAB. First try $s(t)$ corresponding to a square wave or a sawtooth shape, then try something like $s(t) = \cos(0.7t)$.)

k. *Show that if $\vec{x} = \sum_{i=0}^{n-1} \alpha_i \vec{v}_i$ and the $\vec{v}_i$ from an orthonormal set, then using the regular Euclidean norm, $\|\vec{x}\|^2 = \sum_{i=0}^{n-1} |\alpha_i|^2$.*

(HINT: express the norm squared as the inner product with itself and then use the properties of linearity and orthonormality.)

l. *BONUS: Extend (d) to the continuous-time finite duration case. Suppose that you have a continuous real-valued measurement $d(t)$ defined over the interval $t \in [0, T]$ and you want to model it as the weighted sum of m real-valued signals $a_i(t)$ all defined over the same interval. Give and justify an algorithm involving integrals and finite-sized matrix operations that will give the choice of α_i coefficients that minimizes the energy in the error signal $e(t) = d(t) - \sum_{i=0}^{m-1} \alpha_i a_i(t)$.*