Candidacy Exam Presentation
Solutions for Dr. Rhody
Dec 16, 2004

— Transparencies
Let $X$ and $Y$ be Hilbert spaces, and let

$$A : X \rightarrow Y$$
$$B : Y \rightarrow X$$

be bounded linear operators which satisfy

$$\langle Ax, y \rangle = \langle x, By \rangle \quad \forall x \in X \land y \in Y.$$
Hilbert Space: A set $X$ where the metric space is a set $X$ together with a function $d$ (called the "metric" used to measure distance between elements in $X$) such that for every $x, y, z \in X$,

\[ d(x, y) \geq 0 \]
\[ d(x, y) = d(y, x) \]
\[ d(x, y) = 0 \iff x = y \]
\[ d(x, z) \leq d(x, y) + d(y, z) \]

**Hilbert Space**

**Def:** is a vector space $H$ with an inner product $\langle x, y \rangle$ such that the norm defined by

\[ ||x|| = \sqrt{\langle x, x \rangle} \]

turns $H$ into a complete metric space.

**Closed Vector Space:** When applying vector addition, scalar multiplication, or matrix multiplication in a vector plane which results in vectors which are still contained in that plane.
Linear transformations or mapping

\[ A: \mathbf{x} \rightarrow \mathbf{y} \] is a linear transformation from vector space \( \mathbf{x} \) into \( \mathbf{y} \).

Adjoint operators

Let \( A: \mathbf{x} \rightarrow \mathbf{y} \)

The adjoint of the operator \( A \), called \( A^* \), is the operator

\[ A^*: \mathbf{y} \rightarrow \mathbf{x} \]

such that

\[ \langle A^* \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^* \mathbf{y} \rangle \quad \forall \mathbf{x} \in \mathbf{x} \land \mathbf{y} \in \mathbf{y} \]

or

\[ B = A^* \]

so we have

\[ B: \mathbf{y} \rightarrow \mathbf{x} \]

\[ \langle A^* \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, B^* \mathbf{y} \rangle. \]
Adjoint operators: cont.

- The adjoint of a matrix is the conjugate transpose of the matrix,
  \[ A^* = (A^\top) = A^H. \]

- For a real matrix, the adjoint is the transpose,
  \[ A^* = A^\top. \]

- A real matrix that is self-adjoint \((A^* = A)\) is symmetric.

- A complex matrix that is self-adjoint is Hermitian.

- If \(A^*A = AA^*\), the operator is normal.
Range space, \( R(A) \)

Let \( A : X \to Y \)

\[ R(A) = \{ A\vec{x} = \vec{y} : \vec{x} \in X \} \]

- That is, it is the set of values in \( Y \) that are reached from \( X \) by application of \( A \).

\[ R(A) = \text{span}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \]

Set of all lin. comb.

where \( A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n] \).

That is, \( R(A) \) is the space spanned by the columns of matrix \( A \), called the column space.

\[ \vec{y} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \]

\( x_1 \vec{a}_1 + x_2 \vec{a}_2 \ldots \)

- If \( \vec{y} \in R(A) \), then there must be some lin. comb. of the columns of \( A \) that is equal to \( \vec{y} \).

\[ A\vec{x} = \vec{y} \]

Has a solution only if \( \vec{y} \) lies in the col. space of \( A \), i.e. \( \vec{y} \in R(A) \).
Range Space Cont

2) \( \tilde{y} \notin \text{R}(A) \)

There is no solution to \( A\tilde{x} = \tilde{y} \)

3) \( \tilde{y} \in \text{R}(A) \), \( \& \) the cols. of \( A \) are lin. dep.

There are \( \infty \) number of solutions.

Null Space, \( N(A) \)

Let \( A: X \rightarrow Y \)

\[ N(A) = \{ \tilde{x} \in X : A\tilde{x} = \tilde{0} \} \]

- that is, it is the set of vectors in \( X \) that are transformed to \( \tilde{0} \) in \( Y \) by \( A \).

- Furthermore, for \( A\tilde{x} = \tilde{0} \)

1) IF \( \tilde{x} \in N(A) \) and \( \alpha \) is a scalar then,

\[ A(\alpha\tilde{x}) = \alpha A\tilde{x} = \alpha\tilde{0} = \tilde{0} \quad \therefore \alpha\tilde{x} \in N(A). \]

2) IF \( \tilde{x}_p \) is a solution to \( A\tilde{x} = \tilde{y} \) and \( \tilde{x}_n \) is any vector in \( N(A) \), then

\[ \tilde{x} = \tilde{x}_p + \tilde{x}_n \] is also a solution.

\[ A\tilde{x} = A(\tilde{x}_p + \tilde{x}_n) = A\tilde{x}_p + A\tilde{x}_n = \tilde{0} + \tilde{y} = \tilde{y} \]
RANGE of the Adjoint, $R(A^*)$

- another subspace of $A$.
- linear combinations of the conjugates of the rows of $A$.
- $R(A^*)$ is the row space of $A$

(Left)

 Null space of $A^*$, $N(A^*)$

- is the set of vectors $\tilde{\vec{y}}$ such that,$A^H\tilde{\vec{y}} = \vec{0}$ or $\vec{y}^HA = \vec{0}$.

- this is called the left null space.
Orthogonality of spaces

1) - Vectors in the null space of A are orthogonal to all the rows of A.

Before: \[ A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_n \end{bmatrix} = \vec{a}_i \]

Now: \[ A = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_n \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \vec{a}_i^T \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix} = \vec{a}_i \]

Null space defined as: \[ A \vec{x} = \vec{0} \]

If write for each row of A \( \vec{a}_i \), we have \[ \vec{a}_i^T \vec{x} = 0 \] scalars zero

Vectors in \( R(A^*) \) vectors in \( N(A) \)

\[ R(A^*)^\perp = N(A) \] Row spc \perp null spc.

And

\[ R(A)^\perp = N(A^*) \] Col spc \perp left null spc.
(1) Show that the vector $\hat{x}$ of **minimum norm** that satisfies $A\hat{x} = \hat{y}$, is given by

$$x = \mathbb{R} \hat{z}$$

where $\hat{z}$ is any solution of $(AB)\hat{z} = \hat{y}$.

![Diagram](image)

**Ins** - This question is related to least-squares and the **minimum norm**, using adjoints AND orthogonal sub-spaces.

- Consider $A\hat{x} = \hat{y}$ AND:

- Here is the case for minimum norm. Let $\hat{x}_n \in \mathbb{N}(A)$. If $\hat{x}_r$ is a solution of

$$A\hat{x} = \hat{y},$$

then so is $\hat{x}_r + \hat{x}_n$. 
That is, you can add any amount of $X_n$ and arrive at the same solution. This is because any amount, $X_n$, will just get masked out (filtered out) by the operator $A$.

However, any amount of $X_n$ implies that $X$ will grow in norm, which is not what we want. We want min norm.

Therefore, $X_n$ must be in the null space of $A$

$$X_n \in \text{null}(A)$$

Otherwise it will contribute to the solution and change the value of $A\hat{X}_0$.

If this is so, then there must be a vector $\hat{X}$ of $\| \cdot \|_{\min}$ satisfying $A\hat{X} = \hat{y}$.
So, in order to minimize the norm, we want $\hat{x}$ to be entirely in the complement of the nullspace of $A$, which is the rowspace of $A$. That is, the solution must be orthogonal to $N(A)$. Thus

$$\hat{x} \in N(A)^{\perp} = R(B).$$

Since we want $\hat{x}$ to be in the rowspace of $A$, $R(B)$, that implies we take the Hermitian transpose of $A$ and find linear combinations of this basis, to obtain a suitable $\hat{x}$. 
That is, solve
\[ \hat{x} = B \hat{z}, \quad \hat{z} \in \mathbb{Y} \]
for \( \hat{z} \).

You now find linear combinations of the rows of the matrix \( A \). In this way, we are finding the \( x \) that doesn't have any \( x_n \) added to it, thus implying minimum norm.

Therefore, since
\[ A \hat{x} = \hat{y} \]
AND
\[ \hat{x} = B \hat{z} \]
we have
\[ A B \hat{z} = \hat{y}. \]
2. Show that the matrix equation \( AX = \tilde{d} \) has a solution iff \( \tilde{d}^H \tilde{d} = 0 \) for every vector \( \tilde{v} \) such that \( B \tilde{v} = 0 \).

\[ \tilde{d} \] defines the left null space.

ANS

In other words, the equation \( AX = \tilde{d} \) has a solution if and only if \( \langle \tilde{d}, \tilde{v} \rangle = 0 \) for every \( \tilde{v} \in N(B) \).

- You can only get a solution (to \( AX = \tilde{d} \)) if \( \tilde{d} \) lies in the col span of \( A \), \( \tilde{d} \in R(A) \).
- If \( \tilde{d} \notin R(A) \), then there is no solution.

- We know by def. \( R(A) \) is \( \perp N(B) \) and can show that each \( \tilde{v} \in N(B) \) is \( \perp \tilde{d} \).

\[
\langle \tilde{d}, \tilde{v} \rangle = \langle AX, \tilde{v} \rangle = \langle X, B \tilde{v} \rangle = \langle X, 0 \rangle = 0.
\]

Thus

\[ \text{DEF Adjoint: } \langle Ax, y \rangle = \langle x, Ay \rangle \]
Suppose that an equation $A \hat{x} = \hat{d}$ has no $b$ solution. Show that the solution $\hat{x}$ that minimizes $\|A\hat{x} - \hat{d}\|$ satisfies

$$B \hat{x} = B \hat{d},$$

where $\|\cdot\|$ is the induced norm.

**Ans:** If we have no solution, then $\hat{d} \notin \text{R}(A)$.

However, we can find a solution, $\hat{x}$, that minimizes the error, $\hat{e}$, associated with $A\hat{x} - \hat{d}$. That is, minimize $\|A\hat{x} - \hat{d}\|$.

He we wish to derive the stated normal eqn. using an $\hat{x}$ that minimizes $\|A\hat{x} - \hat{d}\|$.

Minimizing $\|A\hat{x} - \hat{d}\|$ is equivalent to minimizing $\|\hat{d} - \hat{d}\|$, where $\hat{d} \notin \text{R}(A)$.

At this point we can take advantage of the projection theorem to help in "minimization".
\( \hat{d} \) is minimum iff \((\hat{d} - \hat{d}) \perp R(A)\).

\[ \hat{d} - \hat{d} \in R(A) \]

or \( \hat{d} - \hat{d} \in N(B) \), since \( R(A) \perp N(B) \),

we can write this in terms of the definition of the left null space \((A^\top y = 0)\),

\[ B(\hat{d} - \hat{d}) = 0 \quad \text{where} \quad A^* = B, \quad y = \hat{d} - \hat{d} \]

or \[ B\hat{d} - B\hat{d} = 0 \]

\[ B\hat{d} = B\hat{d} \]

But \( \hat{d} = A\hat{x} \)

\[ B(A\hat{x}) = B\hat{d} \]

Furthermore,

\[ B(A\hat{x} - \hat{d}) = 0 \quad \text{DEF of} \ N(B) \]

so that \((A\hat{x} - \hat{d}) \) or \( \hat{e} \in N(B) \) which is \( \bot R(A) \) (has a solution)

and has minimal length via the projection thm.
RELATE the above (question 3) to the pseudoinverse of $A$.

\[ A^+ \]

\[ A^+ x = \hat{d} \]

**ANS** - In Q3 we showed that if $A^2 = d$ has no solution ($d \notin R(A)$), we could at least find an $\hat{x}$ that solved $A^\hat{x} = \hat{d}$ in a least-square sense, where $\hat{d} \in R(A)$.

We can explicitly write the normal equation in terms of $\hat{x}$. That is

\[ \hat{x} = (BA)^{-1} B \hat{d} \]

where we see that $(BA)^{-1} B$ is called the pseudoinverse. Formally, the pseudoinverse $A^+$ is defined as the operator that maps $\hat{d}$ to $\hat{x}$ for each $\hat{d} \in Y$. 

\[ \hat{d} = \hat{d}_{\text{ins}} + \hat{d}_{\text{cs}} \]
From the graph we see that $A^+$ takes a point in $Y$ back to a point in $\hat{x} \in \text{R}(B)$ in such a way that $\hat{x}$ has minimum norm (using a projection).

So if $\hat{d} \notin \text{R}(A)$ then we first project $\hat{d}$ back onto $\text{R}(A)$ using the projection $P$. Then map back to $\text{R}(B)$ to a vector $\hat{x}$ of minimum length.

The projection onto $\text{R}(A)$ minimizes the error, $\bar{e}$, or $(\hat{d} - A\hat{x}) = (\hat{d} - \hat{d})$.

\[
\hat{d} = A\hat{x} \\
\hat{x} = (BA)^{-1}B\hat{d}
\]

\[
\hat{d} = A\left[\frac{(BA)^{-1}B\hat{d}}{A^+}\right]
\]

\[
= AA^+\hat{d}
\]

\[
\hat{d} = P\hat{d}
\]
(Additionally)

Candidacy Exam Presentation

Solutions for Dr. Schott

Dec 16, 2004

— Transparencies
PREVIOUS IDEAS

$\alpha_i = \frac{y_i}{e_i}$

for $i = 1, 2, \ldots, N$

Products $N$ abundances.

$\alpha_{ji} = \frac{y_{ij} - a_{ij}}{e_{ij}}$

for $i = 1, 2, \ldots, N$ pixels

for $j = 1, 2, \ldots, K$ target vectors

$\Sigma_{e}$ or $\Sigma_{t}$

$w_i = W_i, w_0 = W_0 = \infty$

$1 - \text{using single target vector \( z \) from image?}$

Matrix of ordered abundances.

Re-ordered

Avg $a_i$:

TARGET Abundance
Hi

Hi

But not a target

Hi

Low INF

Inf
- Pick target vectors that are inside 10, 20, etc.
- Better than hull vectors

- What if z-distribution is not normal?
- Can't use $\frac{z}{\sqrt{a}}$?
- Make log normal?
New Method

\[ \Sigma_t + \Sigma_n \]

↓ MNF using \( \Sigma_n \)

\[ \frac{\Sigma_t}{\Sigma_b} \]

\[ \frac{\Sigma_t}{\Sigma_b} = \frac{P_{b}^{-1}x}{y} \]

\[ S_b = (P_{y/2}^{-1})^T \Sigma_b \frac{P_{y/2}^{-1}} \]

\[ S_t = (P_{y/2}^{-1})^T \Sigma_t \frac{P_{y/2}^{-1}} \]

Want spread in direction \( \perp \hat{y} \)
Computation of Infeasibility

\[ \dot{a}_i = 1, 2 \ldots N - \text{points} \]

\[ j = 1, 2 \ldots M - \text{bands} \]

\[
\dot{\bar{r}}_i = \frac{\dot{r}_i}{a_i} + (1 - \frac{\dot{r}_i}{a_i}) \dot{\bar{r}}_i
\]

\[
\dot{\bar{r}}_i = \frac{\dot{r}_i}{a_i} \bar{S}_a + (1 - \frac{\dot{r}_i}{a_i}) \bar{S}_b
\]

\[
[1 \times 10] = [X \times 10] [20 \times 10]
\]

\[
\text{INF}_i = \frac{1}{N} \frac{\left\| \bar{r}_i \right\|}{\left\| \bar{r}_N \right\|}
\]