EE5120 Linear Algebra: Tutorial 6, July-Dec 2018, Dr. Uday Khankhoje, EE IIT Madras
Covers 5.3, 5.5,5.6 of GS

1. Theory of Eigenvalues and eigenvectors can be used to solve differential equations of multiple variables of the form: $\frac{d\mathbf{u}}{dt} = P\mathbf{u}$ for $\mathbf{u}(t)$, given its initial value $\mathbf{u}(0)$. If $\mathbf{u}(t) = ce^{\lambda t}\mathbf{x}$, then $\frac{d\mathbf{u}}{dt} = c\lambda e^{\lambda t}\mathbf{x}$. We can see that $\mathbf{x}$ is the eigenvector with eigenvalue λ for the matrix P. Also, $\mathbf{u}(0) = c\mathbf{x} \Rightarrow \mathbf{u}(t) = ce^{\lambda t}\mathbf{x}$
So, Any given arbitrary initial value can be expanded in terms of the eigenvectors of P matrix, i.e, $\mathbf{u}(0) = \sum_{i=1}^{n} c_i \mathbf{v}_i$. Then, $\mathbf{u}(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t}\mathbf{v}_i$.
Solve the below differential equation for $\mathbf{u}(t)$
$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \mathbf{u}, \quad with \quad \mathbf{u}(0) = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$
Is the output bounded? If not, for what values of $\mathbf{u}(0)$ will $\mathbf{u}(t)$ be bounded?

Hint: $\mathbf{u}(t)$ should be bounded for bounded $\mathbf{u}(0)$.

Solution: To find eigenvalues: $\det(P - \lambda I) = 0$
$$\Rightarrow (0.5 - \lambda)^2 - 0.5^2 = 0 \Rightarrow \lambda^2 - \lambda = 0 \Rightarrow \lambda = 0, 1.$$
To find eigenvectors: $(P - \lambda I)x = 0$, $Px = 0$
$$\Rightarrow (P - I)x = 0 \Rightarrow \begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\Rightarrow Px = 0 \Rightarrow \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
Solving for c_1, c_2 and then writing $u(0)$ in terms of eigenvectors:
$$\Rightarrow u(0) = \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
Writing for $u(t)$:
$$u(t) = e^{0t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 4e^{1t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
The above expression is not bounded due to the e^t term. On the other hand, if this term has zero coefficient, the output will be bounded. This will happen when:
$$\text{If} \quad u(0) = \alpha \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \text{ then u(t) will be bounded.}$$

2. The matrices A and B are said to be similar if there exists an invertible matrix M such that $A = MBM^{-1}$.
   (a) The *identity transformation* takes every vector to itself: $Tx = x$. Find the corresponding matrix, if both the input and output bases are $\mathbf{v_1} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and $\mathbf{v_2} = \begin{bmatrix} 1 & -1 \end{bmatrix}^T$. How is this matrix related to identity matrix? Are they similar?
   (b) If the transformation T is reflection across the 45 degree line in the plane, find its matrix with respect to the standard basis $\mathbf{e_1} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ and $\mathbf{e_2} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$. Find the corresponding matrix when both the input and output bases are $\mathbf{v_1}$ and $\mathbf{v_2}$ as mentioned in (a). Show that these two matrices are similar by finding the matrix M. Give a geometrical interpretation of M.

Solution:

(a) Since both the input and output bases are same, the matrix corresponding to new bases will still be identity matrix. Note that if the input and output bases were different, the corresponding matrix would not be identity. Since both the matrices are identity, any invertible matrix M will satisfy $I = MIM^{-1}$. Hence they are similar.

(b) Consider the projection matrix, that projects the input onto the 45 degree line. This matrix projects onto the line spanned by $\mathbf{a} = \begin{bmatrix}1 & 1\end{bmatrix}^T$. Hence $P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$.
$P = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$. Recall that the reflection matrix about this 45 degree line can be expressed as $R_1 = 2P - I$. $R_1 = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$. R_1 is the reflection matrix in canonical basis. To find the same matrix in new basis, find where do $\mathbf{v_1}$ and $\mathbf{v_2}$ land on applying the transformation, and express them in new basis. $T(\mathbf{v_1}) = \begin{bmatrix}1 & 0\end{bmatrix}^T$, $T(\mathbf{v_2}) = \begin{bmatrix}0 & -1\end{bmatrix}^T$ (Output is in new basis). These two vectors form the columns of new matrix. Reflection matrix in new basis $R_2 = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$. Consider the matrix M whose columns are $\mathbf{v_1}$ and $\mathbf{v_2}$. It can be verified that $R_1 = MR_2M^{-1}$, hence R_1 and R_2 are similar.

Interpretation of $R_1 = MR_2M^{-1}$*:*The matrix M^{-1} can be interpreted as the matrix which takes a vector in R^2 represented in canonical basis as input and outputs the same vector in R^2 but represented in new basis. R_2 takes the vector represented in new basis and does the transformation then outputs the resultant vector in new basis. The matrix M then takes this vector and represents it in canonical basis. On the whole, MR_2M^{-1} takes a vector represented in canonical basis, and does the transformation then outputs the resultant in canonical basis. This is exactly same as the operation done by R_1.

3. Suppose A is a 3×3 symmetric matrix with eigenvalues 0,1,2.
   (a) What properties can be guaranteed for the corresponding unit eigenvectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$?
   (b) In terms of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ describe the nullspace, left nullspace, rowspace, and columnspace of A.
   (c) Find a vector $\mathbf{x}$ that satisfies $A\mathbf{x} = \mathbf{v} + \mathbf{w}$. Is $\mathbf{x}$ unique?
   (d) Under what conditions on $\mathbf{b}$ does $A\mathbf{x} = \mathbf{b}$ have a solution?
   (e) If $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are the columns of S, what are S^{-1} and $S^{-1}AS$?

Solution:

(a) Since A is symmetric, it's unit eigen vectors are orthonormal.

(b) Rank(A)=number of non-zero eigen values = 2. Since $A = A^T$, the nullspace and leftnullspace of A are same, the rowspace and columnspace of A are same. As we know $A\mathbf{u} = \mathbf{0}$, $\mathcal{N} = \{\alpha\mathbf{u}, \alpha \in R\}$ is the 1-dimensional nullspace and leftnullspace of A. $\mathbf{v}$ and $\mathbf{w}$ will span the rowspace and columnspace.

(c) Consider $\mathbf{x} = \mathbf{v} + \frac{1}{2}\mathbf{w}$. $A\mathbf{x} = \mathbf{v} + \mathbf{w}$. Since A has non trivial nullspace, $\mathbf{x}$ is not unique. $\mathbf{v} + \frac{1}{2}\mathbf{w} + \mathbf{n}$ will also give $\mathbf{v} + \mathbf{w}$ for any $\mathbf{n} \in \mathcal{N}$.

(d) $\mathbf{b}$ should be in coulumnspace of A, i.e $\mathbf{b}$ should be equal to $\alpha\mathbf{v} + \beta\mathbf{w}$ for some $\alpha, \beta \in R$.

(e) Recall the diagonalisation of a symmetric matrix $A = S\Lambda S^{-1} \implies S^{-1}AS = \Lambda$, where the columns of S are unit orthogonal eigenvectors of A and the digaonal entries of the diagonal matrix Λ are the eigen values of A. Here, the diagonal entries of Λ are 0,1 and 2.

4. Let A be an $n \times n$ complex matrix. Assume $A = [\mathbf{a}_1\ \mathbf{a}_2\ ...\ \mathbf{a}_n]$, where $\mathbf{a}_i$ refers to the i^{th} column of matrix A. Define a parameter μ_A, for matrix A, as,

$$\mu_A = \max_{k \neq l} \frac{|\mathbf{a}_k^H \mathbf{a}_l|}{||\mathbf{a}_k||_2 ||\mathbf{a}_l||_2}.$$

In the literature of compressive sensing, μ_A is called the mutual coherence parameter of matrix A. Recall $||\mathbf{a}_i||_2 = \sqrt{\mathbf{a}_i^H \mathbf{a}_i}$, where $\mathbf{a}_i^H$ denotes hermitian (i.e., complex conjugate transpose) of $\mathbf{a}_i$, and define $B = A^H A$.

(a) Show that $0 \leq \mu_A \leq 1$.

(b) Denote $[B]_{i,j}$ as the $(i,j)^{th}$ entry in matrix B. Prove that for every λ being an eigenvalue of B, there exists at least one row, say some m^{th} row, of B such that,

$$\Big|\lambda - [B]_{m,m}\Big| \leq \sum_{p=1, p\neq m}^{n} \Big|[B]_{m,p}\Big|.$$

Refer to the technique used to solve Q9) in tutorial 5 and follow a similar procedure here too. Also, the above result is independent of the information assumed that $B = A^H A$. It is true for any complex square matrix.

(c) Suppose $\mathbf{x}$ is some arbitrary n-length non-zero complex column vector such that $||\mathbf{x}||^2 = 1$, then prove that,

$$\lambda_{\text{min}} \leq \mathbf{x}^H B \mathbf{x} \leq \lambda_{\text{max}},$$

where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of the matrix B. From this part onwards, the information that $B = A^H A$ is needed.

(d) Suppose all the columns of A have unit norm, then prove that λ_{min} and λ_{max}, as defined in (c), are bounded as,

$$\lambda_{\text{max}} \leq 1 + (n-1)\mu_A, \text{ and}$$
$$\lambda_{\text{min}} \geq 1 - (n-1)\mu_A.$$

Make use of the result derived in (b) to obtain the above equations.

(e) Finally, deduce that for a general vector $\mathbf{x} \neq \mathbf{0}$ and matrix A, the generalized result combining (c) and (d) will look like,

$$a_{\text{min}} - a_{\text{max}}(n-1)\mu_A \leq \frac{\mathbf{x}^H B \mathbf{x}}{||\mathbf{x}||^2} \leq a_{\text{max}}\Big(1 + (n-1)\mu_A\Big),$$

where $a_{\text{min}} = \min_{1 \leq k \leq n} ||\mathbf{a}_k||^2$ and $a_{\text{max}} = \max_{1 \leq k \leq n} ||\mathbf{a}_k||^2$.

Solution: Throughout this solution assume m^{th} entry in a vector $\mathbf{y}$ as y_m.

(a) Here, we have,

$$\begin{aligned}
\text{LHS} &= \frac{|\mathbf{a}_k^H \mathbf{a}_l|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} = \frac{\left|\sum_{m=1}^{n} a_{km}^* a_{lm}\right|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} \\
&\leq \sum_{m=1}^{n} \frac{|a_{km}||a_{lm}|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} = \sum_{m=1}^{n} \frac{|a_{km}|}{||\mathbf{a}_k||_2} \frac{|a_{lm}|}{||\mathbf{a}_l||_2} \\
&\leq \sum_{m=1}^{n} \frac{1}{2}\Big(\frac{|a_{km}|^2}{||\mathbf{a}_k||^2} + \frac{|a_{lm}|^2}{||\mathbf{a}_l||^2}\Big) \qquad \Big[\text{Using } (a-b)^2 \geq 0 \Rightarrow ab \leq 0.5(a^2+b^2), \forall a, b \geq 0\Big] \\
&= \frac{\sum_{m=1}^{n} |a_{km}|^2}{2||\mathbf{a}_k||^2} + \frac{\sum_{m=1}^{n} |a_{lm}|^2}{2||\mathbf{a}_l||^2} = \frac{||\mathbf{a}_k||^2}{2||\mathbf{a}_k||^2} + \frac{||\mathbf{a}_l||^2}{2||\mathbf{a}_l||^2} \\
&= 1.
\end{aligned}$$

This is independent of k and l, hence, $\max_{k \neq l} \frac{|\mathbf{a}_k^H \mathbf{a}_l|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} = \mu_A \leq 1$. Further, $|\mathbf{a}_k^H \mathbf{a}_l| \geq 0$, $\forall k, l$. This implies, $\frac{|\mathbf{a}_k^H \mathbf{a}_l|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} \geq 0$, $\forall k, l$. Thus, $\max_{k \neq l} \frac{|\mathbf{a}_k^H \mathbf{a}_l|}{||\mathbf{a}_k||_2||\mathbf{a}_l||_2} \geq 0 \Rightarrow \mu_A \geq 0$.

(b) The result proved in this part is popularly called as **Gershgorin's circle theorem**. B is an $n \times n$ matrix. Suppose λ is some eigenvalue of B and $\mathbf{v}$ be its eigenvector. Define $\mathbf{u} = \frac{\mathbf{v}}{v_s}$, where $s = \arg \max_{1 \leq q \leq n} |v_q|$. Clearly, $\mathbf{u}$ is also an eigenvector of B corresponding to λ and there exists some entry in $\mathbf{u}$ equal to 1. Let it be some m^{th} element, i.e., $u_m = 1$. Then, note that $|u_q| \leq 1$, $\forall q \neq m$. Now, we have, $B\mathbf{u} = \lambda \mathbf{u}$. Concentrating on the m^{th} element on the LHS and RHS vectors, we get,

$$\begin{aligned}
&\sum_{p=1}^{n} [B]_{m,p} u_p = \lambda u_m = \lambda \\
\Rightarrow &\sum_{p=1, p \neq m}^{n} [B]_{m,p} u_p + [B]_{m,m} = \lambda \\
\Rightarrow &\Big| \sum_{p=1, p \neq m}^{n} [B]_{m,p} u_p \Big| = |\lambda - [B]_{m,m}| \\
\Rightarrow &|\lambda - [B]_{m,m}| \leq \sum_{p=1, p \neq m}^{n} |[B]_{m,p}||u_p| \leq \sum_{p=1, p \neq m}^{n} |[B]_{m,p}|(1) \\
\Rightarrow &|\lambda - [B]_{m,m}| \leq \sum_{p=1, p \neq m}^{n} |[B]_{m,p}|.
\end{aligned}$$

Hence, proved.

(c) Since, $B = A^H A$, B is an hermitian matrix. Let its eigenvalue decomposition be $B = U \Lambda U^H$. Now,

$$\mathbf{x}^H B \mathbf{x} = \mathbf{x}^H U \Lambda U^H \mathbf{x} = \mathbf{y}^H \Lambda \mathbf{y}.$$

Let i^{th} diagonal element in the diagonal matrix Λ be λ_i. Hence, we have,

$$\mathbf{y}^H \Lambda \mathbf{y} = \sum_{i=1}^{n} \lambda_i |y_i|^2.$$

Now, $\sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \sum_{i=1}^{n} \lambda_{\max} |y_i|^2 = \lambda_{\max} \sum_{i=1}^{n} |y_i|^2 = \lambda_{\max} ||\mathbf{y}||^2$. Here, $\lambda_{\max}$ is as defined in the question. But, $\mathbf{y} = U^H \mathbf{x} \Rightarrow ||\mathbf{y}||^2 = \mathbf{y}^H \mathbf{y} = \mathbf{x}^H U U^H \mathbf{x} = \mathbf{x}^H \mathbf{x} = ||\mathbf{x}||^2$. Hence, we get,

$$\mathbf{x}^H B \mathbf{x} \leq \lambda_{\max} ||\mathbf{x}||^2. \tag{1}$$

Similarly, $\sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \sum_{i=1}^{n} \lambda_{\min} |y_i|^2 = \lambda_{\min} ||\mathbf{y}||^2 = \lambda_{\min} ||\mathbf{x}||^2$. Also, $\lambda_{\min}$ is as defined in the question. This implies,

$$\mathbf{x}^H B \mathbf{x} \geq \lambda_{\min} ||\mathbf{x}||^2. \tag{2}$$

Combining equations (1) and (2), we get,

$$\lambda_{\min} ||\mathbf{x}||^2 \leq \mathbf{x}^H B \mathbf{x} \leq \lambda_{\max} ||\mathbf{x}||^2. \tag{3}$$

We will be using the above result in proving (e) part. But to arrive at the result for (c), just use the given fact that $\mathbf{x}$ is unit norm, i.e., $||\mathbf{x}||^2 = 1$, in the above equation. Further, the result stated in equation (3) is called **Rayleigh-Ritz theorem**.

(d) Here it is given that A has unit norm columns. So, $||\mathbf{a}_i||^2 = 1$, $\forall i = 1, ..., n$, where $\mathbf{a}_i$ refers to the i^{th} column of matrix A. So, μ_A reduces to, $\mu_A = \max_{k \neq l} |\mathbf{a}_k^H \mathbf{a}_l|$. Also, recall that $B = A^H A$. As a result, $(p, q)^{th}$ entry of B is given by,

$$[B]_{p,q} = \mathbf{a}_p^H \mathbf{a}_q.$$

Hence, if $p = q$, we get $[B]_{p,p} = \mathbf{a}_p^H \mathbf{a}_p = ||\mathbf{a}_p||^2 = 1$, $\forall 1 \leq p \leq n$. And, if $p \neq q$, then,

$$\left| [B]_{p,q} \right| = |\mathbf{a}_p^H \mathbf{a}_q| \leq \mu_A, \tag{4}$$

where we used the definition of μ_A. From (c), we get,

$$\begin{aligned}
\left| \lambda - [B]_{m,m} \right| &\leq \sum_{p=1, p\neq m}^{n} |[B]_{m,p}| \\
\Rightarrow |\lambda - 1| &\leq \sum_{p=1, p\neq m}^{n} |[B]_{m,p}| \leq \sum_{p=1, p\neq m}^{n} \mu_A = (n-1)\mu_A \\
\Rightarrow |\lambda - 1| &\leq (n-1)\mu_A.
\end{aligned}$$

In the above, we made use of equation (4). Clearly, the last equation stated above is independent of the row index of B. Thus, it holds true for all eigen values of B. Now, we get,

$$\begin{aligned}
&|\lambda - 1| \leq (n-1)\mu_A \\
\Rightarrow &-(n-1)\mu_A \leq \lambda - 1 \leq (n-1)\mu_A \\
\Rightarrow &1 - (n-1)\mu_A \leq \lambda \leq 1 + (n-1)\mu_A.
\end{aligned}$$

Since, the above holds true for all eigenvalues we get,

$$\lambda_{\max} \leq 1 + (n-1)\mu_A; \lambda_{\min} \geq 1 - (n-1)\mu_A.$$

(e) Now for a general matrix A without unit norm columns, let $a_{\max} = \max_{1 \le k \le n} ||\mathbf{a}_i||^2$ and $a_{\min} = \min_{1 \le k \le n} ||\mathbf{a}_i||^2$. Then, for all $p \neq q$, the $(p,q)^{th}$ entry in B will be,

$$\left|[B]_{p,q}\right| \le |\mathbf{a}_p^H \mathbf{a}_q| \le a_{\max} \frac{|\mathbf{a}_p^H \mathbf{a}_q|}{\max_{1 \le i \le n} ||\mathbf{a}_i||^2} \le a_{\max} \frac{|\mathbf{a}_p^H \mathbf{a}_q|}{||\mathbf{a}_p||_2 ||\mathbf{a}_q||_2} \le a_{\max} \mu_A.$$

Recall the step in (d),

$$\left|\lambda - [B]_{m,m}\right| \le \sum_{p=1, p \neq m}^{n} |[B]_{m,p}| \le \sum_{p=1, p \neq m}^{n} a_{\max} \mu_A = a_{\max}(n-1)\mu_A.$$

Finally, we get,

$$[B]_{m,m} - a_{\max}(n-1)\mu_A \le \lambda \le [B]_{m,m} + a_{\max}(n-1)\mu_A. \tag{5}$$

Note that $[B]_{m,m} = \mathbf{a}_m^H \mathbf{a}_m = ||\mathbf{a}_m||^2$. So, we have, $a_{\min} \le [B]_{m,m} \le a_{\max}$. Thus, equation (5) can be re-written as,

$$a_{\min} - a_{\max}(n-1)\mu_A \le \lambda \le a_{\max} + a_{\max}(n-1)\mu_A.$$

The above is true for any eigenvalue of B. Therefore, we get,

$$\lambda_{\max} \le a_{\max}\Big(1 + (n-1)\mu_A\Big); \lambda_{\min} \ge a_{\min} - a_{\max}(n-1)\mu_A.$$

Inserting above result in equation (3) derived in (c), we get the desired result.

5. Define matrix D as, $D = [A_1 \ A_2 \ ... \ A_K \ B_1 \ B_2 \ ... \ B_M]$, where all the sub-matrices A_i's and B_j's are of size $m \times n$ with $m > n$ and has unit norm columns. Let μ_D denote the mutual coherence of matrix D (refer to Q 4. for definition of mutual coherence). Suppose,

$$\mathbf{y} = \sum_{k=1}^{K} A_k \mathbf{x}_k,$$

where $\mathbf{x}_k$'s are $n \times 1$ vectors such that $||\mathbf{x}_1||^2 = ||\mathbf{x}_2||^2 = ... = ||\mathbf{x}_K||^2$.

(a) Show that,

$$||A_k^H \mathbf{y}||_2 \ge \Big[1 - (Kn-1)\mu_D\Big] ||\mathbf{x}_k||_2.$$

First, try to lower bound the LHS term as $||P\mathbf{x}_k||_2 - a$, where P is an hermitian matrix and a is an appropriate scalar. For this, you might have to use the following fact: For any two column vectors $\mathbf{u}, \mathbf{v}$, the statement $||\mathbf{u} + \mathbf{v}||_2 \ge ||\mathbf{u}||_2 - ||\mathbf{v}||_2$ holds true. After which, incorporate the results stated as questions in Q 4 to arrive at the desired inequality equation.

(b) Now, prove that the term $||B_m^H \mathbf{y}||_2$ can be upper bounded as,

$$||B_m^H \mathbf{y}||_2 \le Kn\mu_D ||\mathbf{x}_k||_2,$$

for some $k = 1, ..., K$.

(c) As a last step, prove that $\max_{1\le l\le K} ||A_l^H \mathbf{y}||_2 > \max_{1\le m\le M} ||B_m^H \mathbf{y}||_2$ can hold true if,

$$K < \frac{1}{2n}\Big(1+\frac{1}{\mu_D}\Big).$$

The above is an important result obtained in *Greedy Algorithm* theory. The scenario has been simplified in this question to make the entire derivation straight-forward. However, the final result provides a sufficient condition under which a particular greedy algorithm will be able to solve a special type of compressive sensing problem.

Solution: In this solution, suppose Q is a square matrix, we denote $\lambda_{\min}(Q)$ as minimum eigen value of matrix Q and $\lambda_{\max}(Q)$ as it's maximum eigenvalue.

(a) We have,

$$\begin{aligned}||A_k^H \mathbf{y}||_2 &= ||A_k^H\Big(\sum_{l=1}^{K} A_l \mathbf{x}_l\Big)||_2\\ &= ||A_k^H A_k \mathbf{x}_k + \sum_{l=1,l\neq k}^{K} A_k^H A_l \mathbf{x}_l||_2\\ &\ge ||A_k^H A_k \mathbf{x}_k||_2 - ||\sum_{l=1,l\neq k}^{K} A_k^H A_l \mathbf{x}_l||_2.\end{aligned}$$

Now, we will concentrate on each term in the RHS.

$$\begin{aligned}||A_k^H A_k \mathbf{x}_k||_2 &= \sqrt{\mathbf{x}_k^H C_k^H C_k \mathbf{x}_k} \quad [C_k = A_k^H A_k]\\ &\ge \sqrt{\lambda_{\min}(C_k^H C_k)||\mathbf{x}_k||^2} \quad [C_k^H C_k \text{ - is hermitian matrix.}]\\ &= \sqrt{\lambda_{\min}(C_k^2)}||\mathbf{x}_k||_2 \quad [C_k \text{ - is hermitian matrix}]\\ &= \sqrt{\Big(\lambda_{\min}(C_k)\Big)^2}||\mathbf{x}_k||_2\\ &= \lambda_{\min}(C_k)||\mathbf{x}_k||_2 = \lambda_{\min}(A_k^H A_k)||\mathbf{x}_k||_2.\end{aligned}$$

Using the result from Q4(d), we get,

$$\lambda_{\min}(A_k^H A_k) \ge 1-(n-1)\mu_D,$$

where n is the row (and column) size of the hermitian matrix $A_k^H A_k$. Thus, we obtain,

$$||A_k^H A_k \mathbf{x}_k||_2 \ge \lambda_{\min}(A_k^H A_k)||\mathbf{x}_k||_2 \ge \Big[1-(n-1)\mu_D\Big]||\mathbf{x}_k||_2. \tag{6}$$

Now, we focus on the term $||\sum_{l=1,l\neq k}^{K} A_k^H A_l \mathbf{x}_l||_2$.

$$\begin{aligned}||\sum_{l=1,l\neq k}^{K} A_k^H A_l \mathbf{x}_l||_2 &\le \sum_{l=1,l\neq k}^{K} ||A_k^H A_l \mathbf{x}_l||_2\\ &= \sum_{l=1,l\neq k}^{K} \sqrt{\mathbf{x}_l^H A_l^H A_k A_k^H A_l \mathbf{x}_l}\\ &\sum_{l=1,l\neq k}^{K} \sqrt{\lambda_{\max}(E)||\mathbf{x}_l||_2},\end{aligned}$$

where $E = A_l^H A_k A_k^H A_l$ and clearly E is an hermitian matrix of size $n \times n$. Let $\tilde{E} = A_k^H A_l$. Note that we are considering the case $l \neq k$. So, the $(i,j)^{th}$ entry in $\tilde{E}$ is s.t.,

$$\left|[\tilde{E}]_{i,j}\right| = |\mathbf{a}_{k_i}^H \mathbf{a}_{l_j}|,$$

where $\mathbf{a}_{k_i}$ and $\mathbf{a}_{l_j}$ refer to i^{th} and j^{th} columns of matrices A_k and A_l respectively. So, $\left|[\tilde{E}]_{i,j}\right| = |\mathbf{a}_{k_i}^H \mathbf{a}_{l_j}| \leq \mu_D$. Further, $(i,j)^{th}$ entry of E is such that,

$$\begin{aligned}
\left|[E]_{i,j}\right| &= \left|\tilde{\mathbf{e}}_i^H \tilde{\mathbf{e}}_j\right| \\
&= \left|\sum_{p=1}^{n} [\tilde{E}]_{n,i}^* [\tilde{E}]_{n,j}\right| \\
&\leq \sum_{p=1}^{n} \left|[\tilde{E}]_{n,i}\right| \left|[\tilde{E}]_{n,j}\right| \\
&\leq \sum_{p=1}^{n} \mu_D^2 \\
&= n\mu_D^2.
\end{aligned}$$

Now, using the procedure adopted in Q4(d), we get,

$$\lambda_{\max}(E) \leq n\mu_D^2 + (n-1)n\mu_D^2 = n^2\mu_D^2.$$

Thus, we get,

$$||\sum_{l=1, l\neq k}^{K} A_k^H A_l \mathbf{x}_l||_2 \leq \sum_{l=1, l\neq k}^{K} \sqrt{n^2\mu_D^2}||\mathbf{x}_l||_2 = (K-1)n\mu_D||\mathbf{x}_k||_2. \quad (7)$$

Combining equations (6) and (7), we finally get a lower bound on $||A_k^H \mathbf{y}||_2$ as,

$$\begin{aligned}
||A_k^H \mathbf{y}||_2 &\geq \left[1 - (n-1)\mu_D - (K-1)n\mu_D\right]||\mathbf{x}_k||_2 \\
&= \left[1 - (Kn-1)\mu_D\right]||\mathbf{x}_k||_2.
\end{aligned}$$

Hence, proved.

(b) In this case, we have,

$$\begin{aligned}
||B_m^H \mathbf{y}||_2 &= ||B_m^H \sum_{k=1}^{K} A_k \mathbf{x}_k||_2 \\
&\leq \sum_{k=1}^{K} ||B_m^H A_k \mathbf{x}_H||_2 \\
&\leq \sum_{k=1}^{K} \sqrt{\lambda_{\max}(A_k^H B_m B_m^H A_k)}||\mathbf{x}_k||_2 \\
&\leq \sum_{k=1}^{K} n\mu_D||\mathbf{x}_k||_2 = Kn\mu_D||\mathbf{x}_k||_2.
\end{aligned}$$

Here, the upper bound for $\lambda_{\max}(A_k^H B_m B_m^H A_k)$ is obtained in exactly the same way as that obtained for $\lambda_{\max}(A_l^H A_k A_k^H A_l)$ when $k \neq l$.

(c) We can see that $\max_{1\le k\le K} ||A_k^H \mathbf{y}||_2 \ge \Big[1-(Kn-1)\mu_D\Big]||\mathbf{x}_k||_2$. Also, $\max_{1\le m\le M} ||B_m^H \mathbf{y}||_2 \le Kn\mu_D||\mathbf{x}_k||_2$. The condition,

$$\max_{1\le k\le K} ||A_k^H \mathbf{y}||_2 > \max_{1\le m\le M} ||B_m^H \mathbf{y}||_2,$$

will be satisfied if the lower bound of the LHS term is greater than the upper bound of the RHS term. Hence, we get,

$$\Big[1-(Kn-1)\mu_D\Big]||\mathbf{x}_k||_2 > Kn\mu_D||\mathbf{x}_k||_2.$$

On re-arranging the terms we can arrive at the desired result.

6. Consider two adjoining cells separated by a permeable membrane, and suppose that a fluid flows from the first cell to the second one at a rate (in milliliters per minute) that is numerically equal to three times the volume (in milliliters) of the fluid in the first cell. It then flows out of the second cell at a rate (in milliliters per minute) that is numerically equal to twice the volume in the second cell. Let $x_1(t)$ and $x_2(t)$ denote the volumes of the fluid in the first and second cells at time t, respectively. Assume that, initially, the first cell has 40 milliliters of fluid, while the second one has 5 milliliters of fluid. Find the volume of fluid in each cell at time t.

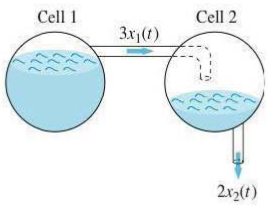

Figure 1: Figure for Q.6

Solution: No fluid is flows into the first cell and fluid flows from the first cell to the second cell is equal to three times the volume of the fluid in the first cell. So,

$$\frac{dx_1(t)}{dt} = -3x_1(t)$$

The change in volume of the fluid in the second cell is given by

$$\frac{dx_2(t)}{dt} = 3x_1(t) - 2x_2(t)$$

This can be written in matrix form as

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

The eigenvalues of the matrix

$$A = \begin{bmatrix} -3 & 0 \\ 3 & -2 \end{bmatrix}$$

are $\lambda_1 = -3$, and $\lambda_2 = -2$ and corresponding eigen vectors are

$$\begin{bmatrix} 1 \\ -3 \end{bmatrix} \textit{and} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Hence the general solution is given by

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ -3 \end{bmatrix} e^{-3t} + b_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-2t}$$

Using the initial conditions,
$b_1 = 40$, and $b_2 = 125$
Thus the volume of fluid in each cell at time t is given by

$$x_1(t) = 40e^{-3t}$$
$$x_2(t) = -120e^{-3t} + 125e^{-2t}$$

7. (a) Prove that $C(A), N(A), C(A^H)$ and $N(A^H)$ are the fundamental spaces in complex case, i.e, give their properties and derive them. Verify for the matrix

$$A = \begin{bmatrix} 1 & i & 0 \\ i & 0 & 1 \end{bmatrix}$$

(b) Prove that determinant of a Hermitian matrix is real.

Solution:

(a) Let A is a matrix of size $m \times n$.

- Proof that $N(A)$ is orthogonal to $C(A^H)$.
i.e, We should show that $v^H w = 0$, if $w \in C(A^H)$ and $v \in N(A)$.

$$w \in C(A^H) \Rightarrow \exists\, u \in \mathbb{C}^n \quad s.t \quad w = A^H u$$
$$\Rightarrow v^H w = v^H A^H u = (Av)^H u = (0)^H u$$
$$\Rightarrow v^H w = 0 \quad \forall w \in C(A^H)$$

So, $N(A)$ is orthogonal to $C(A^H)$.
- Proof that $N(A^H)$ is orthogonal to $C(A)$.
i.e, We should show that $v^H w = 0$, if $w \in C(A)$ and $v \in N(A^H)$.

$$w \in C(A) \Rightarrow \exists\, u \in \mathbb{C}^n \quad s.t \quad w = Au$$
$$\Rightarrow v^H w = v^H A u = (A^H v)^H u = (0)^H u = 0$$
$$\Rightarrow v^H w = 0 \quad \forall w \in C(A)$$

So, $N(A^H)$ is orthogonal to $C(A)$.
- Proof that $C(A), C(A^H), N(A), N(A^H)$ are subspaces.
Let $w_1, w_2 \in C(A) \Rightarrow \exists u_1, u_2 \in \mathbb{C}^n \quad s.t \quad Au_1 = w_1,\ Au_2 = w_2$
$\Rightarrow A(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 w_1 + \alpha_2 w_2 \Rightarrow \alpha_1 w_1 + \alpha_2 w_2 \in C(A)$
and $A(0_{n\times 1}) = 0_{m\times 1} \Rightarrow 0_{m\times 1} \in C(A)$, so C(A) is a subspace.
Similarly do it for $C(A^H), N(A), N(A^H)$.

- Let $dim(N(A)) = r$ and $u_1, u_2, .., u_r$ be the basis of $N(A)$.
Since $u_1, u_2, ..., u_r$ is a linearly independent set in $\mathbb{C}^n$, we can extend it to form a basis of $\mathbb{C}^n$. Now there exists vectors $u_{r+1}, u_{r+2}, ..., u_n$ such that the set $\{u_1, \ldots, u_r, u_{r+1}, ..., u_n\}$ is a basis of $\mathbb{C}^n$.
Span of $C(A) =$ span of $\{Au_1, .., Au_r, Au_{r+1}, ..Au_n\}$ = span of $\{0, .., 0, Au_{r+1}, ..Au_n\}$
$\Rightarrow$ span of $C(A) =$ span of $\{Au_{r+1}, ..Au_n\}$.
we now prove that $\{Au_{r+1}, ..Au_n\}$ are linearly independent. Suppose the set is linearly dependent, then there exist scalars
$alpha_{r+1}, ..\alpha_n\}$ not all zero such that $\alpha_{r+1}Au_{r+1} + ... + \alpha_n Au_n = 0$.

$$\Rightarrow A(\alpha_{r+1}u_{r+1} + ... + \alpha_n u_n) = 0$$

This implies $\alpha_{r+1}u_{r+1} + ... + \alpha_n u_n$ belongs to null space of A.

$$\Rightarrow \alpha_{r+1}u_{r+1} + ... + \alpha_n u_n = \alpha_1 u_1 + ... + \alpha_r u_r$$

But, as $u_1, ..., u_n$ are linearly independent only possibility is all α's are zero.
So, $\{Au_{r+1}, ..Au_n\}$ are linearly independent.
So, $dim(C(A)) = n - r$.
i.e, $dim(C(A)) + dim(N(A)) = n$
Similarly prove $dim(C(A^H)) + dim(N(A^H)) = m$

Row reduced echolon form of $A = \begin{bmatrix} 1 & i & 0 \\ i & 0 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & -i \\ 0 & 1 & 1 \end{bmatrix}$.

So, $C(A) : \begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} i \\ 0 \end{bmatrix}$ and $N(A) : \begin{bmatrix} i \\ -1 \\ 1 \end{bmatrix}$

Row reduced echolon form of $A^H = \begin{bmatrix} 1 & -i \\ -i & 0 \\ 0 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

So, $C(A^H) : \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}, \begin{bmatrix} -i \\ 0 \\ 1 \end{bmatrix}$ and $N(A^H) : \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

(b) We can see this for 1×1 matrix directly. i.e, $det(c^H) = c^H = conj(c) = conj(det(c))$, where c is a scalar.
We know that, $det(A) = det(A^T) \Rightarrow det(A^H) = det(A^{H^T})$
i.e, A^{H^T} is each element of A replaced by its complex conjugate.
If you see the expansion of $det(A^{H^T})$, each term is same as $det(A)$ except elements replaced by their complex conjugates. Taking the conjugate out.

$$\Rightarrow det(A^{H^T}) = det(conj(A)) = conj(det(A))$$

Now, If A is Hermitian:

$$A^H = A \Rightarrow det(A^H) = det(A)$$

But, we have seen that $det(A^H) = conj(det(A))$

$$\Rightarrow det(A^H) = conj(det(A)) = det(A) \Rightarrow det(A) = conj(det(A))$$

So, $det(A)$ is real.

8. The Harmonic Oscillator is shown in figure 2, where a particle is constrained by some kind of forces like in the case of atoms in solids. Find the energy values that can be taken by particle by solving the following eigen equation:

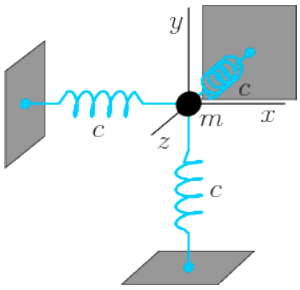

Figure 2: Classical picture of an harmonic oscillator

$$\mathbf{H}\psi = E\psi$$

where ψ is the eigen function which gives the probability of finding the particle at (x,y,z), E is the corresponding eigen energy value and $\mathbf{H}$ is the Hamiltonian given as below,

$$\mathbf{H} = -\frac{\hbar^2}{2m}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) + \frac{c}{2}(x^2 + y^2 + z^2)$$

Hint: For other eigenfunctions take $\psi_l = exp(\alpha l^2)f(l)$. $f(l)$ will be Hermite polynomials.
Hint: Take $\psi_l = exp(-\alpha l^2)$ for first eigen function. As probability of finding the particle at ∞ is 0.
Hint: Take $\psi(x,y,z) = \psi_x(x)\psi_y(y)\psi_z(z)$. Eigen functions are continuous equivalent of eigenvectors.

Solution: Using Variable separable method, write $\psi(x,y,z) = \psi_x(x)\psi_y(y)\psi_z(z)$ and substitute in eigen equation.

$$\mathbf{H}\psi_x(x)\psi_y(y)\psi_z(z) = E\psi_x(x)\psi_y(y)\psi_z(z)$$

$$[-\frac{\hbar^2}{2m}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) + \frac{c}{2}(x^2 + y^2 + z^2)]\psi_x(x)\psi_y(y)\psi_z(z) = E\psi_x(x)\psi_y(y)\psi_z(z)$$

On doing partial derivatives and then dividing by ψ:

$$[-\frac{\hbar^2}{2m}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) + \frac{c}{2}(x^2 + y^2 + z^2)]\psi_x(x)\psi_y(y)\psi_z(z) = E\psi_x(x)\psi_y(y)\psi_z(z)$$

$$\Rightarrow -\frac{\hbar^2}{2m\psi_x(x)}\frac{\partial^2\psi_x(x)}{\partial x^2} + \frac{cx^2}{2} + -\frac{\hbar^2}{2m\psi_y(y)}\frac{\partial^2\psi_y(y)}{\partial y^2} + \frac{cy^2}{2} + -\frac{\hbar^2}{2m\psi_z(z)}\frac{\partial^2\psi_z(z)}{\partial z^2} + \frac{cz^2}{2} = E$$

writing the constant $E = E_x + E_y + E_z$, We can solve the following equation:

$$-\frac{\hbar^2}{2m\psi_l(l)}\frac{\partial^2\psi_l(l)}{\partial l^2} + \frac{cl^2}{2} = E_l \quad \forall l = x, y, z$$

One of the solutions will be of the form ae^{-bl^2}, where a is a normalization constant used to see that total probability of finding the particle anywhere from $(-\infty, \infty)$ is 1.

$$\Rightarrow -\frac{\hbar^2}{2mae^{-bl^2}}a[-2be^{-bl^2} + 4b^2l^2e^{-bl^2}] + \frac{cl^2}{2} = E_l$$

$$\Rightarrow -\frac{\hbar^2}{2m}[-2b + 4b^2l^2] + \frac{cl^2}{2} = E_l$$

$$\Rightarrow [2b\frac{\hbar^2}{2m} - E_l] + l^2[\frac{c}{2} - 4b^2\frac{\hbar^2}{2m}] = 0$$

This has to satisfy for all l,

$$\Rightarrow b = \pm(\frac{cm}{4\hbar^2})^{1/2}, \quad E_l = \pm(\frac{cm}{4\hbar^2})^{1/2}\frac{\hbar^2}{m}$$

We take only $E_{l_0} = \frac{\hbar}{2}(\frac{c}{m})^{1/2}$, as in the other case there will be finite probability for particle to stay at ∞, which is not physical as the force is trying to restrict to some space. And corresponding eigen function as $\psi_l 0 = a_0 \exp[-(\frac{cm}{4\hbar^2})^{1/2}l^2]$. Similar to the above expression, we expect other eigen functions also to have zero at $\pm\infty$. So, We take $\psi_l = f(l)e^{-kl^2}$ and substitute in eigen equation to get,

$$\frac{\partial^2 f(l)}{\partial l^2} - 2k\frac{\partial f(l)}{\partial l} + (k^2 - 1) - \frac{2m}{\hbar^2}(\frac{c}{2} - E_l) = 0$$

When we write $f(l) = \sum_{i=0}^{\infty} p_i l^p$, and solve it will be observed that p_i has a recursive relation. As we need ψ_l to be finite everywhere, importantly when $l \to \pm\infty$. Only eigenvalues possible are :

$$E_{l_n} = (n_l + \frac{1}{2})\frac{\hbar}{2}(\frac{c}{m})^{1/2}$$

$$\Rightarrow E = E_x + E_y + E_z = (n_x + n_y + n_z + \frac{3}{2})\frac{\hbar}{2}(\frac{c}{m})^{1/2}$$

That is energy levels that are allowed are quantized.

Refer to : `https://quantummechanics.ucsd.edu/ph130a/130_notes/node153.html`

The main take away is:

- Eigen functions are continuous equivalents of eigenvectors.
- Also, In case you have no analytic solution possible and domain is fintie, like if we say that particle can only exist between (-L,L), then you can discretize the problem into finite points and get to a matrix form as $A\psi = \lambda\psi$. (Solve the problem computationally).