18.06 Problem Set 8 - Solutions

Due Wednesday, 14 November 2007 at 4 pm in 2-106.

Problem 1: (20=5+5+5+5) Consider the matrix $A = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}$.

(a) Check that A is a positive Markov matrix, and find its steady state.

$\boxed{\text{Solution}}$ A is obviously positive Markov matrix, since all entries are positive, and each column sum is 1.

Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is a corresponding eigenvector, then

$$\begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

which implies $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Thus the steady state is $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$.

(b) Factor A into $S\Lambda S^{-1}$.

$\boxed{\text{Solution}}$ Since $\mathrm{tr}(A) = 1.5$ and $\lambda_1 = 1$, we see $\lambda_2 = 0.5$. The corresponding eigenvector $\mathbf{v}_2$ satisfies $\begin{pmatrix} 0.3 & 0.3 \\ 0.2 & 0.2 \end{pmatrix} \mathbf{v}_2 = 0$, i.e. $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. So the eigenvector matrix is $S = \begin{pmatrix} 0.6 & -1 \\ 0.4 & 1 \end{pmatrix}$, whose inverse is $S^{-1} = \begin{pmatrix} 1 & 1 \\ -0.4 & 0.6 \end{pmatrix}$. So

$$A = \begin{pmatrix} 0.6 & -1 \\ 0.4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -0.4 & 0.6 \end{pmatrix}.$$

(Remark: since we can take multiples of the above vectors as eigenvectors, the decomposition above is not unique.)

(c) Explain why A^k approaches $A^\infty = \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix}$ in two ways, using results in (a) and (b) respectively.

$\boxed{\text{Solution}}$

Method 1: we have

$$A^\infty = A^\infty A = \begin{pmatrix} A^\infty \mathbf{u}_0 & A^\infty \mathbf{v}_0 \end{pmatrix},$$

where $\mathbf{u}_0$ and $\mathbf{v}_0$ are two columns of A. From part (a) we see that $A^\infty \mathbf{u}_0$ is some multiple of $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$. Since the sum of entries is conserved, we see $A^\infty \mathbf{u}_0 = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$. By the same reason, $A^\infty \mathbf{v}_0 = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$. This implies the conclusion.

Method 2: From part (b) we get

$$\begin{aligned} A^\infty &= \begin{pmatrix} 0.6 & -1 \\ 0.4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix}^\infty \begin{pmatrix} 1 & 1 \\ -0.4 & 0.6 \end{pmatrix} \\ &= \begin{pmatrix} 0.6 & -1 \\ 0.4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -0.4 & 0.6 \end{pmatrix} \\ &= \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix} \end{aligned}$$

(d) Find all Markov matrices with steady state $(0.6, 0.4)^T$.

$\boxed{\text{Solution}}$ Such Markov matrix has to be a 2×2 matrix. Let $A = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}$. Since $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$ is the steady state, we have $\begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix} \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$. This is equivalent to the relation $3a + 2b = 3$, i.e. $b = 3(1-a)/2$.

Moreover, since the entries of A are nonnegative, a and b should satisfies $0 \le a \le 1$ and $0 \le b \le 1$. Notice that the condition $b \le 1$ together with $3a+2b = 3$ implies $a \ge 1/3$.

We conclude that all possible Markov matrices are

$$A = \begin{pmatrix} a & 3(1-a)/2 \\ 1-a & (3a-1)/2 \end{pmatrix}$$

with $1/3 \le a \le 1$.

(However, we may study the two possible cases in more detail. First case $a = 1/3$ and $b = 1$, then $A = \begin{pmatrix} 1/3 & 1 \\ 2/3 & 0 \end{pmatrix}$. By the trace-trick, the second eigenvalue is $-2/3$, so $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$ is the only steady state. The second case $a = 1$ and $b = 0$, then A is the identity matrix, and thus the vector $\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$ is not the only steady state – vectors in other direction will not tends to this vector.)

Problem 2: (10) Suppose A is a Markov matrix, and let $\mathbf{y} = A\mathbf{x}$ for some vector $\mathbf{x}$. Show that the sum of the components of $\mathbf{y}$ equals the sum of the components of $\mathbf{x}$ — e.g. if the components of $\mathbf{x}$ are populations, then A conserves the total population. (Hint: recall the proof from class that the product of two Markov matrices is a Markov matrix.)

Solution For any vector $\mathbf{x}$, the sum of the components of $\mathbf{x}$ equals

$$x_1 + \cdots + x_n = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \mathbf{x}.$$

Since A is Markov, we have

$$\begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} A = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}.$$

Thus the sum of components

$$\begin{aligned} y_1 + \cdots + y_n &= \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \mathbf{y} \\ &= \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} A\mathbf{x} \\ &= \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \mathbf{x} \\ &= x_1 + \cdots + x_n. \end{aligned}$$

This completes the proof.

(Alternately, we can derive this using the explicit summations: If $\mathbf{y} = A\mathbf{x}$, then $y_i = \sum_j A_{ij} x_j$. Therefore, $\sum_i y_i = \sum_i \sum_j A_{ij} x_j = \sum_j (\sum_i A_{ij}) x_j = \sum_j x_j$, where we have used the fact that $\sum_i A_{ij} = 1$ (the sum of each column of A is 1).)

Problem 3: (10=4+4+2) In class we learned that any positive Markov matrix A has a dominant eigenvalue, λ (=1), in the sense that it is simple eigenvalue (not a repeated root) and larger than the absolute value of any other eigenvalues. As a consequence, we know that any vector $\mathbf{x}$ will approach a multiple of the corresponding eigenvector $\mathbf{v}_1$ when we apply A to $\mathbf{x}$ again and again. In fact, this property holds for general matrix with positive entries, no matter whether is Markov or not.

(a) Use MATLAB to construct a random 5×5 positive matrix (`A = rand(5,5)` will give you a positive matrix), and use `[S,D]=eig(A)` to find its eigenvalues `diag(D)` and corresponding eigenvectors (columns of S). What is the dominant eigenvalue? Do the same procedure three times more, with 6×6, 7×7, and 8×8 random positive matrices respectively.

Solution The codes are

```
A=rand(5,5);[S,D]=eig(A)

S =

   0.5755   0.3457 - 0.2719i   0.3457 + 0.2719i  -0.4957   0.4757
   0.5121  -0.3048 + 0.0845i  -0.3048 - 0.0845i  -0.0453  -0.0138
   0.3288   0.1307 - 0.1722i   0.1307 + 0.1722i  -0.1410  -0.1041
   0.4415   0.6317             0.6317            -0.3699   0.5518
   0.3218  -0.4331 + 0.2706i  -0.4331 - 0.2706i   0.7717  -0.6769


D =

   2.3041        0                  0                  0        0
        0   0.2403 + 0.1937i        0                  0        0
        0        0             0.2403 - 0.1937i        0        0
        0        0                  0            -0.0617        0
        0        0                  0                  0   0.1118

A=rand(6,6);[S,D]=eig(A)

S =

   0.4547    0.0769   -0.2412 - 0.2755i  -0.2412 + 0.2755i  -0.3190   -0.1879
   0.4628   -0.3439   -0.0085 + 0.3875i  -0.0085 - 0.3875i   0.0527   -0.3743
   0.4954   -0.0246   -0.2876 + 0.0544i  -0.2876 - 0.0544i   0.5464    0.4606
   0.2991   -0.5746    0.2079 - 0.4206i   0.2079 + 0.4206i  -0.5158   -0.1764
   0.3583    0.7239    0.4969             0.4969            -0.2297   -0.2782
   0.3403   -0.1453   -0.3392 + 0.2188i  -0.3392 - 0.2188i   0.5273    0.7098


D =

   3.2675         0         0                  0                  0         0
        0   -0.7348         0                  0                  0         0
        0         0   -0.0009 + 0.3287i        0                  0         0
        0         0         0            -0.0009 - 0.3287i        0         0
        0         0         0                  0             0.1725         0
```

```
       0         0         0                 0                 0    0.5764

A=rand(7,7);[S,D]=eig(A)

S =

-0.4451 -0.3344 + 0.0239i  -0.3344 - 0.0239i  -0.6160   0.0878 + 0.0769i  0.0878 - 0.076
-0.3524 -0.0673 - 0.1015i  -0.0673 + 0.1015i   0.0740   0.0364 + 0.0131i  0.0364 - 0.013
-0.5108  0.1873 - 0.0095i   0.1873 + 0.0095i   0.2407  -0.3057 + 0.0321i -0.3057 - 0.032
-0.2858  0.7892             0.7892            -0.3947   0.2644 - 0.0712i  0.2644 + 0.071
-0.3892 -0.0890 + 0.0509i  -0.0890 - 0.0509i   0.5000   0.2878 + 0.0817i  0.2878 - 0.081
-0.2648 -0.3517 + 0.0398i  -0.3517 - 0.0398i   0.2664   0.1875 - 0.1188i  0.1875 + 0.118
-0.3370 -0.2809 - 0.0009i  -0.2809 + 0.0009i   0.2834  -0.8226           -0.8226


D =

   3.3641         0            0            0        0             0            0
       0   -0.7064 + 0.0703i  0            0        0             0            0
       0         0      -0.7064 - 0.0703i  0        0             0            0
       0         0            0       0.0501        0             0            0
       0         0            0            0   0.4928 + 0.1129i   0            0
       0         0            0            0        0       0.4928 - 0.1129i   0
       0         0            0            0        0             0        0.6919

A=rand(8,8);[S,D]=eig(A)

S =

-0.3534  0.5837   0.0787 - 0.2610i   0.0787 + 0.2610i   0.3609   0.0912   0.2868  -0.164
-0.3662  0.1682  -0.0200 + 0.1409i  -0.0200 - 0.1409i  -0.4086  -0.3904  -0.4819   0.352
-0.3146  0.3837  -0.4216 + 0.1647i  -0.4216 - 0.1647i  -0.3610   0.3770   0.1151   0.177
-0.3627  0.0875   0.1859 + 0.3200i   0.1859 - 0.3200i   0.3539  -0.0284  -0.1625   0.316
-0.4438 -0.0533   0.1738 + 0.0969i   0.1738 - 0.0969i   0.1886   0.1441   0.5046  -0.587
-0.3570 -0.3712   0.5113             0.5113            -0.5415   0.7096   0.5003  -0.575
-0.2792 -0.2615  -0.3619 - 0.0767i  -0.3619 + 0.0767i   0.3389  -0.2537  -0.3141   0.151
-0.3288 -0.5168   0.0320 - 0.3531i   0.0320 + 0.3531i  -0.0601  -0.3280  -0.2055   0.136
```

```
D =

   3.9113         0         0                  0                  0         0         0         0
        0   -0.8971         0                  0                  0         0         0         0
        0         0    0.1782 + 0.7242   0                  0         0         0         0
        0         0         0             0.1782 - 0.7242i    0         0         0         0
        0         0         0                  0             0.7315         0         0         0
        0         0         0                  0                  0    0.3784         0         0
        0         0         0                  0                  0         0   -0.0402         0
        0         0         0                  0                  0         0         0   -0.1832
```

The dominant eigenvalues are the first eigenvalues, 2.3041, 3.2675, 3.3641 and 3.9113. The first columns of S'es are the dominant eigenvectors.

(b) You should notice that the dominant eigenvector has components that are all of the same sign (and hence they could all be chosen positive). (This generalizes the property, from class, that the dominant eigenvector of a Markov matrix has components that can be chosen nonnegative.) Prove that this is true in general. (Hint: the dominant eigenvector, when multiplied by A, should grow faster than any other vector. Show that if the dominant eigenvector had two components with different signs, that you could construct a different vector that grows faster when multiplied by A.)

$\boxed{\text{Solution}}$ **WARN: The hint is wrong!** Because the eigenvectors are not necessarily orthogonal, it turns out that the dominant eigenvector does **NOT** necessarily grow faster than any other vector, or more explicitly the dominant eigenvector does not maximize $|Ax|^2/|x|^2$ (rather, this is maximized by the dominant eigenvector of A^TA as a consequence of the minimax principle, which we haven't covered yet. –Thanks Ben.

(To grader: proofs based on the incorrect hint will be accepted.)

Correct proof: Suppose that the dominant eigenvector has differently signed components. Now take a random positive vector and multiply it by A^n for large n. Eventually, the vector must tend to a multiple of the dominant eigenvector, which means that it must eventually have components of different signs. But this is impossible because the vector started out positive and we multiplied it by a positive matrix A—there is no way to ever get a negative number. Thus the dominant eigenvector cannot have differently signed components. (This glosses over one point—how do you know that you can construct a positive vector with a component of the dominant eigenvector? If it were *not* possible, then the span of all the $n-1$ other eigenvectors must include *all* positive vectors. But this is impossible since the set of all positive vectors [which includes n independent vectors, e.g. the columns of I] is not contained in an $n-1$ dimensional subspace.)

(c) Is there any consistent pattern in the signs of the entries of all other eigenvectors?

$\boxed{\text{Solution}}$ The other eigenvectors may be real vector or complex vector. However, even if the eigenvector is real, the components must have different signs. In other words, for a positive matrix, the dominant eigenvector is the only eigenvector which have all entries the same sign!

Problem 4: (20=4+4+3+3+3+3) The purpose of this problem is to re-examine some of the things we did earlier in the course, and to see what changes when we allow the possibility of complex vectors $\mathbf{x}$ and matrices A and the adjoint A^H. Justify your answers.

(a) For a complex A, is the left nullspace $N(A^T)$ orthogonal to $C(A)$ under the old unconjugated inner product $\mathbf{x}^T\mathbf{y}$ or the new conjugated inner product $\mathbf{x}^H\mathbf{y}$? What about $N(A^H)$ and $C(A)$?

$\boxed{\text{Solution}}$ The left nullspace $N(A^T)$ is orthogonal to $C(A)$ under the old unconjugated inner product. In fact, if $\mathbf{u} \in N(A^T)$ and $A\mathbf{v} \in C(A)$, then

$$(A\mathbf{v})^T\mathbf{u} = \mathbf{v}^T(A^T\mathbf{u}) = 0.$$

However, if we use the new conjugated inner product, then the left nullspace $N(A^T)$ is not orthogonal to $C(A)$ in general. For example, take $A = \begin{pmatrix} 1 & 1 \\ i & i \end{pmatrix}$, then $\mathbf{u} = \begin{pmatrix} 1 \\ i \end{pmatrix} \in C(A)$ and we also have $\mathbf{u} \in N(A^T)$, but $\mathbf{u}^H\mathbf{u} = 2 \neq 0$.

On the other hand, the left nullspace $N(A^H)$ is orthogonal to $C(A)$ under the new conjugated inner product: In fact, if $\mathbf{u} \in N(A^H)$ and $A\mathbf{v} \in C(A)$, then

$$(A\mathbf{v})^H\mathbf{u} = \mathbf{v}^H(A^H\mathbf{u}) = 0.$$

Similarly we know that in general $N(A^H)$ is not orthogonal to $C(A)$ under the old unconjugated inner product.

(b) For a real vector subspace, V, the intersection of V and $V^\perp$ is only the single point $\mathbf{0}$. Now suppose V is a complex vector subspace. If we define $V^\perp$ as the set of vectors $\mathbf{x}$ with $\mathbf{x}^T\mathbf{v} = 0$ for all $\mathbf{v} \in V$, give an example of a V that intersects $V^\perp$ at a non-zero vector (hint: the simplest example is probably a 1-dimensional subspace of $\mathbb{C}^2$). What about if we use $\mathbf{x}^H\mathbf{v} = 0$, does V ever intersect $V^\perp$ at a nonzero vector using the conjugated definition of orthogonality? (Mathematicians use the latter definition of $V^\perp$.)

$\boxed{\text{Solution}}$ An example: Let V be the 1 dimensional complex vector subspace of $\mathbb{C}^2$ given by the span of the vector $(1, i)$. Then since $(1,i)^T(1,i) = 0$, we see that $V^\perp = V$, thus $V \cap V^\perp = V$!

If we use $\mathbf{x}^H\mathbf{v} = 0$ to define the orthogonal complement $V^\perp$, then $V \cap V^\perp = \{0\}$. In fact, suppose $V \cap V^\perp$ contains some nonzero vector $\mathbf{x}$, then $\mathbf{x}$ is perpendicular to itself under this conjugated definition of orthogonality, i.e. $\mathbf{x}^H\mathbf{x} = 0$. But $\mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2$. So $\mathbf{x}$ has to be the zero vector.

(c) Using your answer to (b), find an example of an $m \times n$ complex matrix A (for any m and n you like) such that $C(A) + N(A^T) \neq \mathbb{C}^m$, unlike for real matrices where $C(A) + N(A^T) = \mathbb{R}^m$ always (because the two subspaces only intersected at $\mathbf{0}$).

$\boxed{\text{Solution}}$ Take $A = \begin{pmatrix} 1 \\ i \end{pmatrix}$, then $C(A) = N(A^T)$ is the complex vector subspace V spanned by the vector $\begin{pmatrix} 1 \\ i \end{pmatrix}$. Thus $C(A) + N(A^T) = V \neq \mathbb{C}^2$.

(d) Based on (a), (b), and (c), what would you suggest that we use as the four fundamental subspaces for complex matrices?

$\boxed{\text{Solution}}$ We should use the column space $C(A)$, the nullspace $N(A)$, and their orthogonal complements $N(A^H)$ and $C(A^H)$.

(e) We know A and A^T have equal rank. What about A and $\bar{A}$? A and A^H?

$\boxed{\text{Solution}}$ A and $\bar{A}$ also have equal rank. In fact, if λ is an eigenvalue of A with eigenvector $\mathbf{v}$, i.e. $A\mathbf{v} = \lambda\mathbf{v}$, then by taking conjugate we get $\bar{A}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. In other words, $\bar{\lambda}$ is an eigenvalue of $\bar{A}$. This implies that the number of nonzero eigenvalues of A equals the number of nonzero eigenvalues of $\bar{A}$. So they have the same rank.

Since $\text{rank}(A) = \text{rank}(\bar{A})$ and $\text{rank}(\bar{A}) = \text{rank}(\bar{A}^T)$, we see $\text{rank}(A) = \text{rank}(A^H)$.

Alternately: if $U = EA$ is the result of elimination on A to get $\text{rank}(A)$ pivots, the we can complex-conjugate everything to get $\bar{U} = \bar{E}\bar{A}$ — that is, the elimination steps are simply conjugated, and we get the same number of pivots.

(f) How is $\det A$ related to $\det A^H$?

$\boxed{\text{Solution}}$ We have proved above that the eigenvalues of $\bar{A}$ are exactly the conjugate of the eigenvalues of A. Thus $\det \bar{A} = \overline{\det A}$. So $\det A^H = \det(A^H)^T = \det(\bar{A}) = \overline{\det A}$.

Alternately, we could also use the explicit formula (the BIG formula) for the determinant, from which it is obvious that conjugating the matrix conjugates the determinant.

Problem 5: (10=2+2+3+3) Justify the following true statements:

(a) If A is unitary, then A is invertible and A^{-1} is unitary.

$\boxed{\text{Solution}}$ If A is unitary, then $AA^H = I$, thus A is invertible with $A^{-1} = A^H$.

Moreover, since $A^{-1}(A^{-1})^H = A^H(A^H)^H = A^H A = I$, A^{-1} is unitary.

(b) If A and B are unitary, then their product AB is unitary.

$\boxed{\text{Solution}}$ We have $AA^H = BB^H = I$. Thus

$$(AB)(AB)^H = A(BB^H)A^H = AIA^H = AA^H = I.$$

(c) If A is Hermitian and A is invertible, then A^{-1} is also Hermitian.

$\boxed{\text{Solution}}$ Since $AA^{-1} = I$, taking conjugate we get $\bar{A}\overline{A^{-1}} = I$. Taking transpose, we get $(A^{-1})^H A^H = I$. But $A^H = A$, so $(A^{-1})^H A = I$, i.e. $(A^{-1})^H = A^{-1}$. This shows A^{-1} is Hermitian.

(Another proof: we have seen from pset 7 that according to Cayley-Hamilton theorem, A^{-1} is a polynomial of A, thus if A is Hermitian, A^{-1} is automatically Hermitian.)

(d) If A is diagonalizable, then e^A is diagonalizable.

$\boxed{\text{Solution}}$ A is diagonalizable if and only it has n linearly independent eigenvectors $\mathbf{v}_1, \cdots, \mathbf{v}_n$. Denote the corresponding eigenvalues by λ_i. Then we have

$$e^A v_i = (I + A + \frac{1}{2!}A^2 + \cdots)v_i = (1 + \lambda_i + \frac{1}{2!}\lambda_i^2 + \cdots)v_i = e^{\lambda_i} v_i,$$

which implies that v_i is an eigenvector of e^A corresponding to eigenvalue λ_i. So e^A has n linearly independent eigenvectors and thus diagonalizable.

Problem 6: (10=3+4+3) Suppose A is anti-Hermitian, i.e., $A^H = -A$. (This is also called "skew-Hermitian.") Note that a special case of this is a real anti-symmetric matrix, i.e. $A^T = -A$ for real A.)

(a) Show that iA is Hermitian. Conclude that the eigenvalues of A purely imaginary.

$\boxed{\text{Solution}}$ Since $A^H = -A$, we have

$$(iA)^H = \bar{i}A^H = -i(-A) = iA.$$

So the matrix iA is Hermitian. Since the eigenvalues of the Hermitian matrix iA are real, the eigenvalues of the matrix A are purely imaginary numbers (and 0).

(b) Show that e^A is a unitary matrix.

$\boxed{\text{Solution}}$ Since $(A^k)^H = (A^H)^k$ for all k, we have

$$(e^A)^H = e^{(A^H)} = e^{-A}.$$

Since A and $-A$ commutes, we get

$$(e^A)^H e^A = e^{-A}e^A = e^{-A+A} = e^{0I} = I.$$

Thus e^A is unitary.

(c) Show that the solution $\mathbf{u}(t)$ of the system $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ satisfies $\|\mathbf{u}(t)\|^2 = \|\mathbf{u}(0)\|^2$.

$\boxed{\text{Solution}}$ The solution to the system above is $u(t) = e^{At}\mathbf{u}(0)$. Thus

$$\|\mathbf{u}(t)\|^2 = \|e^{At}\mathbf{u}(0)\|^2 = \mathbf{u}(0)^H(e^{At})^H e^{At}\mathbf{u}(0) = \mathbf{u}(0)^H\mathbf{u}(0) = \|\mathbf{u}(0)\|^2.$$

Problem 7: (10) Unlike the exponential function on numbers, the matrix e^Ae^B is in general different from e^Be^A, and both can be different from e^{A+B}. Check the above statement for $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$\boxed{\text{Solution}}$ The characteristic equation of A is $\lambda^2 - 2 = 0$, so $\lambda_1 = \sqrt{2}, \lambda_2 = -\sqrt{2}$.

For $\lambda_1 = \sqrt{2}$, we have $\mathbf{v}_1 = \begin{pmatrix} 1 \\ \sqrt{2}-1 \end{pmatrix}$. For $\lambda_2 = -\sqrt{2}$, we have $\mathbf{v}_2 = \begin{pmatrix} 1-\sqrt{2} \\ 1 \end{pmatrix}$. So the matrix of eigenvectors is $S = \begin{pmatrix} 1 & 1-\sqrt{2} \\ \sqrt{2}-1 & 1 \end{pmatrix}$. So

$$\begin{aligned} e^A &= \begin{pmatrix} 1 & 1-\sqrt{2} \\ \sqrt{2}-1 & 1 \end{pmatrix} \begin{pmatrix} e^{\sqrt{2}} & 0 \\ 0 & e^{-\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 1-\sqrt{2} \\ \sqrt{2}-1 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{4-2\sqrt{2}} \begin{pmatrix} e^{\sqrt{2}} + (\sqrt{2}-1)^2 e^{-\sqrt{2}} & (\sqrt{2}-1)(e^{\sqrt{2}} - e^{-\sqrt{2}}) \\ (\sqrt{2}-1)(e^{\sqrt{2}} - e^{-\sqrt{2}}) & (\sqrt{2}-1)^2 e^{\sqrt{2}} + e^{-\sqrt{2}} \end{pmatrix} \end{aligned}$$

Similarly the matrix B has characteristic equation $\lambda_2 + 1 = 0$, i.e., $\lambda_1 = i, \lambda_2 = -i$.

For $\lambda_1 = i$, we have $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. For $\lambda_2 = -i$, we have $\mathbf{v}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$. So the matrix of eigenvectors is $S = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$. So

$$\begin{aligned} e^B &= \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{pmatrix} \end{aligned}$$

Finally $A + B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$, so $\lambda_1 = 1, \lambda_2 = -1$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. So

$$\begin{aligned} e^{A+B} &= \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}^T \\ &= \begin{pmatrix} e & 0 \\ e - e^{-1} & e^{-1} \end{pmatrix}. \end{aligned}$$

The fact that $e^A e^B$, $e^B e^A$ and e^{A+B} are all different follows from direct computation.

Problem 8: (10) Solve the ODE system $\frac{d\mathbf{u}}{dt} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{u}$ for $\mathbf{u}|_{t=0} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$\boxed{\text{Solution}}$ From solution of problem 7 above we have seen that for $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$,

$$e^{-A} = \begin{pmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{pmatrix}.$$

By the same argument, one can easily see that

$$e^{-At} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Thus

$$e^{At} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

So

$$\mathbf{u}(t) = e^{At} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}$$