Question 3 [15 marks in total] Let A be an $n \times n$ matrix.

1. [2 marks] (Straight forward) Show that $I - A$ is an invertible matrix if and only if 1 is not an eigenvalue of A.

2. [5 marks] (Straight forward) Show that if $I - A$ is invertible, then

$$I + A + A^2 + ... + A^m = (I - A)^{-1}(I - A^{m+1}), \text{ for } m = 0, 1, 2, ...$$

3. [8 marks] (Straight forward) For every $v \in \mathbb{R}^n$, define

$$f_m(v) = v^T v + v^T A v + v^T A^2 v + ... + v^T A^m v = \sum_{j=0}^{m} v^T A^j v,$$

where "v^T" denotes the transpose of v as usual. Assume that A is diagonalizable in complex numbers and all its complex eigenvalues have absolute values (moduli) less than 1. Show that for every $v \in \mathbb{R}^n$, $f_m(v)$ has a limit when $m \to \infty$ and find that limit.

Question 4 [40 marks in total] This question involves the concept of independent random variables; you do not need to know the precise definition below.

Definition 1. *Let $X : \Omega \to \mathbb{R}^m$ and $Y : \Omega \to \mathbb{R}^n$ be two vector-valued random variables on the same probability space $(\Omega, \mathcal{F}, P)$. Then X and Y are called independent of for any rectangles $I \subset \mathbb{R}^m$ and $J \subset \mathbb{R}^n$, $P(X^{-1}(I) \cap Y^{-1}(J)) = P(X^{-1}(I))P(Y^{-1}(J))$.*

We only need the following property of independent random variables: if X and Y are two independent $\mathbb{R}^n$-valued random variables, which means that both X and Y are (measurable) functions from Ω to $\mathbb{R}^n$, and both X and Y have covariance matrices, then the covariance matrix of $X + Y$ is the sum of the covariance matrices of X and Y.

If $Y : \Omega \to \mathbb{R}^n$ is an $\mathbb{R}^n$-valued random variable and A is an $m \times n$ matrix, then we can define an $\mathbb{R}^m$-valued random variable AY as follows:

$$(AY)(\omega) = AY(\omega), \text{ for every } \omega \in \Omega. \tag{1}$$

The left hand side is the value of our new function AY on ω and the right hand side is the product of the $m \times n$ matrix A and $n \times 1$ matrix $Y(\omega)$. A useful fact is that if the covariance matrix of Y is W_Y, then the covariance matrix of AY is AW_YA^T. This fact may be used in this question without proof.

Now consider the following difference equation:

$$Y_{t+1} = AY_t + U_{t+1}, \text{ for } t = 0, 1, 2, ... \tag{2}$$

Here Y_t and U_{t+1} are both $\mathbb{R}^n$-valued random variables for every $t \geq 0$ and A is an $n \times n$ matrix assumed to be known. The AY_t on the right hand side is the new $\mathbb{R}^n$-valued random variable defined in Eq. (1) and the right hand side of Eq. (2) is a sum of two $\mathbb{R}^n$-valued random variables AY_t and U_{t+1}.

Throughout this question, we maintain the following assumptions.

(a) For every $t \geq 0$, Y_t and U_{t+1} are independent random variables. This implies that AY_t and U_{t+1} are also independent for every $t \geq 0$.

(b) The covariance matrix of Y_0 is W_0 and the covariance matrix of U_{t+1} is V for every $t \geq 0$.

In Parts 2 and 3 of this question, we also assume that for every non-zero $\beta \in \mathbb{R}^n$, $\beta^T V \beta > 0$.

1. [10 marks] (Medium) Show that for every $Y_t \geq 0$, Y_t has the covariance matrix

$$W_t = \sum_{j=0}^{t-1} A^j V (A^T)^j + A^t W_0 (A^T)^t. \tag{3}$$

(When $t = 0$, it is agreed that the first term on the right hand side is 0.) Parts 2 and 3 are about properties of the matrix W_t; if you cannot solve this part, you may answer Parts 2 and 3 treating Eq. (3) as the definition of W_t without worrying about where this W_t comes from.

2. [10 marks] (Challenging) A key question in time series analysis is whether W_t converges as $t \to \infty$. Throughout the question we focus on the case where A is diagonalizable in complex numbers: $A = PDP^{-1}$ for some invertible complex matrix P and diagonal matrix D. The diagonal entries of D are denoted by λ_1, λ_2, ..., λ_n. (If you are not familiar with complex numbers, you may assume that A is diagonalizable in real numbers in answering this and the next part and will not be penalized for doing that.) In this part, assume that $|\lambda_j| \leq |\lambda_1|$ for every j and $|\lambda_1| < 1$. Show that there exists an $M > 0$ and $b \in (0, 1)$ such that every matrix entry of $A^j V (A^T)^j$ has absolute value less than Mb^j for every $j \geq 0$. This fact will imply that W_t converges as $t \to \infty$ under our assumptions on the λ_j's, but you do not need to prove this convergence yourself.

3. [10 marks] (Medium) The process Eq. (2) is said to possess a *unit root* if 1 is an eigenvalue of A (or A^T since A and A^T have the same set of eigenvalues). In this part, assume that Eq. (2) does possess a unit root. Show that there exists a non-zero vector $\beta \in \mathbb{R}^n$ such that $\beta^T W_t \beta \to \infty$ as $t \to \infty$.

4. [10 marks] (Challenging) Consider the following process:

$$x_{t+1} = a_1 x_t + a_2 x_{t-1} + a_3 x_{t-2} + \epsilon_{t+1}, \text{ for } t = 2, 3, 4, \ldots. \tag{4}$$

Here x_t and ϵ_t are random variables for $t \geq 0$ and ϵ_{t+1} is independent of (x_t, x_{t-1}, x_{t-2}) for every $t \geq 2$. As Eq. (4) looks very different from Eq. (2), the theory from Parts 1 to 3 of the question do not immediately apply to the process in Eq. (4). Eq. (4) may be rewritten as the following system of equations:

$$\begin{aligned} x_{t+1} &= a_1 x_t + a_2 x_{t-1} + a_3 x_{t-2} + \epsilon_{t+1}; \\ x_t &= x_t + 0 x_{t-1} + 0 x_{t-2} + 0; \\ x_{t-1} &= 0 x_t + x_{t-1} + 0 x_{t-2} + 0. \end{aligned}$$

The first equation is Eq. (4), and the last two equations hold trivially for every $t \geq 2$. Using this trick, convert Eq. (4) to the form in Eq. (2) with appropriately chosen A, Y_t and U_{t+1}. Finally, find the characteristic polynomial of your A in terms of a_1, a_2 and a_3.