EE363 homework 6 solutions

1. *Constant norm and constant speed systems.* The linear dynamical system $\dot{x} = Ax$ is called *constant norm* if for every trajectory x, $\|x(t)\|$ is constant, *i.e.*, doesn't depend on t. The system is called *constant speed* if for every trajectory x, $\|\dot{x}(t)\|$ is constant, *i.e.*, doesn't depend on t.

   (a) Find the (general) conditions on A under which the system is constant norm.

   (b) Find the (general) conditions on A under which the system is constant speed.

   (c) Is every constant norm system a constant speed system?

   (d) Is every constant speed system a constant norm system?

   Solution.
   The system is constant norm if and only if

$$\begin{aligned} 0 &= \frac{d}{dt}\|x(t)\|^2 \\ &= 2x(t)^T\dot{x}(t) \\ &= 2x(t)^T Ax(t) \\ &= x(t)^T(A + A^T)x(t) \end{aligned}$$

   for all $x(t)$, which occurs if and only $A + A^T = 0$, which is the same as $A^T = -A$, *i.e.*, A is skew-symmetric. There are many other ways to see this. For example, the norm of the state will be constant provided the velocity vector is always orthogonal to the position vector, *i.e.*, $\dot{x}(t)^T x(t) = 0$. This also leads us to $A + A^T = 0$.

   Another approach uses the state transition matrix e^{tA}. The system is constant norm provided e^{tA} is orthogonal for all $t \geq 0$. From here, you'd have to argue that A must be skew-symmetric.

   The system is constant speed if and only if

$$\begin{aligned} 0 &= \frac{d}{dt}\|\dot{x}(t)\|^2 \\ &= \frac{d}{dt}\|Ax(t)\|^2 \\ &= 2(Ax(t))^T A\dot{x}(t) \\ &= 2x(t)^T A^T A^2 x(t) \\ &= x(t)^T A^T(A + A^T)Ax(t) \end{aligned}$$

   for all $x(t)$, which occurs if and only $A^T(A + A^T)A = 0$. In other words, the matrix $A^T A^2$ is skew-symmetric.

We see that if a system is constant norm, then it must be constant speed, since $A+A^T = 0$ implies that $A^T(A + A^T)A = 0$.

But the converse is false, as the simple system

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x,$$

which is a double integrator, shows. This system has trajectories of the form

$$x(t) = \begin{bmatrix} x_1(0) + tx_2(0) \\ x_2(0) \end{bmatrix}.$$

It doesn't have constant norm, but it does have constant speed, since $\dot{x} = (x_2(0), 0)$.

2. *An iterative method for solving the ARE.* We consider the LQR problem with linear system $\dot{x} = Ax + Bu$, and state and input weight matrices Q and R, with $Q = Q^T > 0$ and $R = R^T > 0$. (The positive definiteness assumption on Q is made for convenience only; implies that (Q, A) is observable.) The optimal input has the form $u(t) = Kx(t)$, where $K = -R^{-1}B^TP$, where P is the unique positive definite solution of the ARE

$$A^TP + PA + Q - PBR^{-1}B^TP = 0.$$

   (a) Show that the ARE is equivalent to the two equations

   $$(A + BK)^TP + P(A + BK) + (Q + K^TRK) = 0, \qquad K = -R^{-1}B^TP.$$

   The first equation is a Lyapunov equation (in P).

   (b) Let K_0 be any matrix for which $A + BK_0$ is stable. Let P_0 the solution of the Lyapunov equation above, with $K = K_0$. From P_0, we define $K_1 = -R^{-1}B^TP_0$. Now repeat, *i.e.*, let P_1 be solution of the Lyapunov equation above with $K = K_1$, and so on. Show that $A + BK_i$ are all stable, and P_i are all positive definite.

   Hint. Use induction. To show that $A + BK_i$ is stable, show that

   $$(A+BK_i)^TP_{i-1}+P_{i-1}(A+BK_i)+\Big(Q + K_i^TRK_i + (K_i - K_{i-1})^TR(K_i - K_{i-1})\Big) = 0,$$

   and use a Lyapunov theorem.

   (c) Show that $P_{i+1} \leq P_i$. The sequence $P_1, P_2, \ldots$ is nonincreasing and bounded below by 0, so it converges to some limit P. Show that this limit is the solution of ARE.

   (d) Run the algorithm on a numerical example. You can choose A stable, for example as `A=randn(10); A=A-1.1*max(real(eig(A)))*eye(10);`, and use $K_0 = 0$. Plot $\|P_i - P\|$, where P is the solution of the ARE. If you've got it right, a small number of steps (say, 10) should more than suffice.

Solution:

(a) Substituting $K = -R^{-1}B^T P$ into

$$(A+BK)^T P + P(A+BK) + (Q + K^T RK) = 0$$

gives us

$$(A - BR^{-1}B^T P)^T P + P(A - BR^{-1}B^T P) + (Q + PBR^{-1}B^T P) = 0,$$

which simplifies to the ARE

$$A^T P + PA + Q - PBR^{-1}B^T P = 0.$$

(b) Subtracting $(A+BK_i)^T P_{i-1} + P_{i-1}(A+BK_i) + (Q + K_i^T RK_i)$ from both sides of the Lyapunov equation for P_{i-1},

$$(A+BK_{i-1})^T P_{i-1} + P_{i-1}(A+BK_{i-1}) + Q + K_{i-1}^T RK_{i-1} = 0$$

we get

$$\begin{aligned}(A+BK_i)^T P_{i-1} + P_{i-1}(A+BK_i) + (Q + K_i^T RK_i) = \\ (K_i - K_{i-1})^T B^T P_{i-1} + P_{i-1}B(K_i - K_{i-1}) + K_i^T RK_i - K_{i-1}^T RK_{i-1}.\end{aligned}$$

Since $K_i = -R^{-1}B^T P_{i-1}$, we have $B^T P_{i-1} = -RK_i$. Substituting this into the above equation gives

$$(A+BK_i)^T P_{i-1} + P_{i-1}(A+BK_i) + \Big(Q + K_i^T RK_i + (K_i - K_{i-1})^T R(K_i - K_{i-1})\Big) = 0.$$

Now suppose that P_{i-1} is positive definite. Since $A + BK_i$ satisfies the above Lyapunov equation, and $Q + K_i^T RK_i + (K_i - K_{i-1})^T R(K_i - K_{i-1}) > 0$, $A + BK_i$ must be stable. If $A + BK_i$ is stable, the solution to the Lyapunov equation

$$(A+BK_i)^T P_i + P_i(A+BK_i) + Q + K_i^T RK_i = 0,$$

is positive definite, since $Q + K_i^T RK_i > 0$. Thus we have shown that if $P_{i-1} > 0$, then $A + BK_i$ is stable, and $P_i > 0$. Since $A + BK_0$ is stable, P_0 must be positive definite. So by induction, $A + BK_i$ are all stable, and P_i are all positive definite.

(c) Subtracting the Lyapunov equation

$$(A+BK_i)^T P_i + P_i(A+BK_i) + (Q + K_i^T RK_i) = 0$$

from

$$(A+BK_i)^T P_{i-1} + P_{i-1}(A+BK_i) + \Big(Q + K_i^T RK_i + (K_i - K_{i-1})^T R(K_i - K_{i-1})\Big) = 0$$

we get

$$(A+BK_i)^T(P_{i-1}-P_i)+(P_{i-1}-P_i)(A+BK_i)+(K_i-K_{i-1})^TR(K_i-K_{i-1})=0.$$

Since $A+BK_i$ is stable and $(K_i-K_{i-1})^TR(K_i-K_{i-1})\geq 0$, we conclude that $P_{i-1}\geq P_i$.

To show that this converges to the solution of the ARE, note that $K_i=-R^{-1}B^TP_{i-1}$, and so the control gains converge to $K=-R^{-1}B^TP$, where P is the limit of the sequence $P_1,P_2,\ldots$. We know that P must satisfy the equation

$$(A+BK)^TP+P(A+BK)+(Q+K^TRK)=0.$$

Substituting $K=-R^{-1}B^TP$ we get

$$A^TP+PA-PBR^{-1}B^TP=0,$$

which shows that P solves the ARE.

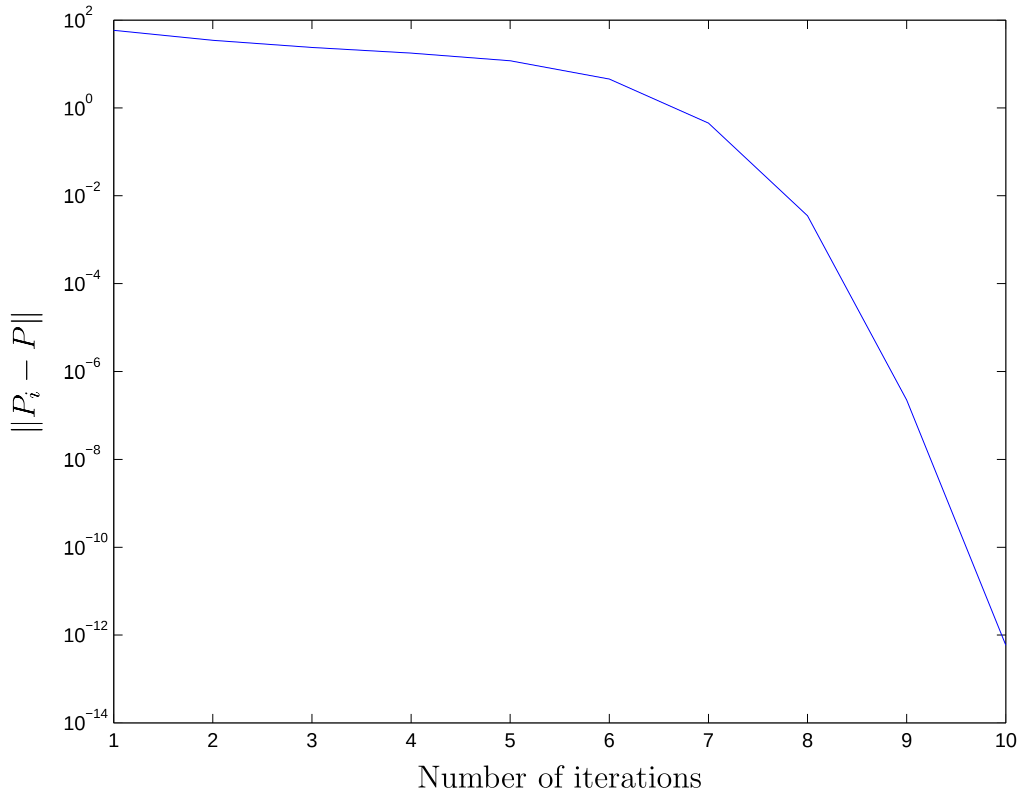
(d) The following code implements this method.

```
randn('state',0);
n = 10; m = 3;
A = randn(n); B = randn(n,m);
A = A-1.1*max(real(eig(A)))*eye(n);
R = randn(m); R = R'*R;
Q = randn(n); Q = Q'*Q;
K0 = zeros(m,n);
Pare = are(A,B*inv(R)*B',Q);

nsteps = 10; K = K0; J = zeros(nsteps,1);
for i = 1:10
    P = lyap((A+B*K)',Q+K'*R*K);
    K = -inv(R)*(B'*P);
    J(i) = norm(P-Pare);
end

semilogy(J); xlabel('n'); ylabel('J');
print('-depsc','kleinman_rock.eps');
```

The following figure plots $\|P_i-P\|$ (notice the quadratic convergence).



3. *A Lyapunov condition for attraction.* A set $C \subseteq \mathbf{R}^n$ is said to be *attractive* or an *attractor* for $\dot{x} = f(x)$, if every trajectory eventually ends up in (and stays in) C. More precisely, for any trajectory x, there is a time T (which can depend on the trajectory) such that $x(t) \in C$ for $t \geq T$.

   Note the subtle difference between an invariant set and an attractor. If a trajectory enters an invariant set, it will stay in the set thereafter. For an attractor set, *every* trajectory eventually enters (and then stays).

   Establish the following Lyapunov attractor theorem: Suppose there is a function $V : \mathbf{R}^n \to \mathbf{R}$, and constants $a > 0$ and b such that for all z,

$$V(z) \geq b \Longrightarrow \dot{V}(z) \leq -a.$$

   Then the set $C = \{z \mid V(z) \leq b\}$ is an attractor.

   Solution: We start by showing that the sublevel set $V(z) \leq b$ is invariant. Consider the boundary *i.e.* $\mathcal{C} = \{z | V(z) = b\}$, for every $z \in \mathcal{C}$ we have $\dot{V}(z) < -a$ where a is a positive constant. So $V(z)$ is decreasing on the boundary, hence for every $x(t)$ on the boundary $x(t + \delta t)$ will be inside $\mathcal{C}$, allowing us to conclude that $\mathcal{C}$ is an invariant set.

   Now we need to show that all trajectories will eventually wind up in $\mathcal{C}$, we recall that for all z, $\dot{V}(z) < -a$ giving

$$V(x(t)) = V(x(0)) + \int_{\tau=0}^{t} \dot{V}(x(\tau))d\tau < V(x(0)) + \int_{\tau=0}^{t} -a d\tau = V(x(0)) - at.$$

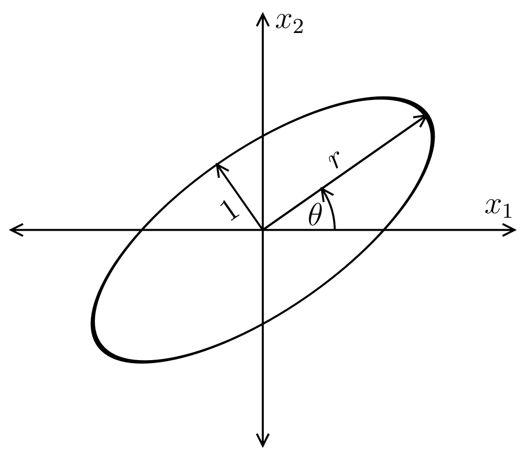
$V(x(0))$ is a constant for all trajectories and $a > 0$, so we can always find a T such that $V(x(0)) - aT < b$ hence $V(x(T)) < b$ and therefore $x(T) \in \mathcal{C}$, and the set $\mathcal{C}$ is indeed an *attractor*.

4. *Finding an invariant ellipsoid for a linear system.* Consider the linear system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 4 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Recall that an ellipsoid $\mathcal{E}$ is said to be invariant for this system if all trajectories that start in $\mathcal{E}$ stay in $\mathcal{E}$, *i.e.*, $x(0) \in \mathcal{E}$ implies that $x(t) \in \mathcal{E}$ for all $t \geq 0$.

You will find an invariant ellipsoid for this system. We will describe the ellipsoid by the length $r \geq 1$ of its major semi-axis (the length of the minor semi-axis is set to one) and the angle θ of the major semi-axis with respect to the x_1-axis, as shown in the figure below.



(a) First present a general description of how you will go about finding r and θ, briefly justifying each step.

(b) Carry out the individual steps in your description from part (a) to find specific values of r and θ.

Solution:

(a) The solution to this problem relies on two facts:

- The ellipsoid $\mathcal{E}$ shown in the figure can be expressed as

$$\mathcal{E} = \{z \mid z^T P z \leq 1\}$$

for some $P > 0$.

- The set $\{z \mid z^T P z \leq 1\}$ is invariant for the linear system $\dot{x} = Ax$ if and only if $A^T P + PA \leq 0$.

In fact, we can express P explicitly in terms of r and θ:

$$P = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1/r^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

The easiest thing to do is to first find some $P > 0$ such that $A^T P + PA \leq 0$, and then normalize P so that the largest eigenvalue of P is one, and then solve for r and θ.

So, a general description is as follows:

- *Find $P > 0$ such that $A^T P + PA \leq 0$.* There are several ways of doing this. One way is to solve the Lyapunov equation

$$A^T P + PA + Q = 0,$$

  where $Q > 0$ can be arbitrarily chosen, *e.g.*, $Q = I$.
- *Normalize.* We first diagonalize P, and then normalize P by dividing it by its larger eigenvalue.
- *Extract r and θ.* The square-root of the reciprocal of the smaller eigenvalue of the normalized P gives r. θ can be obtained by computing the angle between an eigenvector corresponding to the smaller eigenvalue of P and the x-axis.

(b) First we will find P by solving $A^T P + PA + I = 0$. We get

$$P = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & \frac{9}{2} \end{bmatrix}$$

The eigenvalues of P are $(5 + \sqrt{20})/2$ and $(5 - \sqrt{20})/2$. Therefore,

$$r = \left(\frac{5 + \sqrt{20}}{5 - \sqrt{20}} \right)^{1/2} \approx 4.24.$$

Next compute θ. First we find that

$$\begin{bmatrix} \frac{1}{2} & 1 \\ 1 & \frac{9}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 - \sqrt{5} \end{bmatrix} = \frac{5 - \sqrt{20}}{2} \begin{bmatrix} 1 \\ 2 - \sqrt{5} \end{bmatrix}$$

Therefore θ is the angle between the x-axis and $(1, 2, -\sqrt{5})$, which is just

$$\theta = \tan^{-1}\left(2 - \sqrt{5}\right) \approx -0.23 \text{ radians}.$$

5. *Global asymptotic stability for a system with small nonlinearity.* We consider the system $\dot{x} = Ax + q(x)$, where $x(t) \in \mathbf{R}^n$ and $q : \mathbf{R}^n \to \mathbf{R}^n$. We assume that q satisfies $\|q(z)\| \leq \alpha \|z\|$ for all z, but otherwise is unknown. We assume that A is stable, *i.e.*, all its eigenvalues have negative real part.

Intuition suggests that for α small, the system is globally asymptotically stable. Show that this intuition is correct, by finding a positive number $\bar{\alpha}$ such that for $\alpha \leq \bar{\alpha}$, you can guarantee global asymptotic stability of the system. The number $\bar{\alpha}$ must be explicit, *i.e.*, it should be easily calculated using standard matrix operations (such as computing eigenvalues and singular values, solving Lyapunov or Riccati equations, etc.), and the problem data.

Solution: We first solve the Lyapunov equation $A^T P + PA + I = 0$ to get P. (We can do this because A is stable, so the Lyapunov operator fs nonsingular.) We also know that $P > 0$. Let's use the quadratic Lyapunov function $V(z) = z^T Pz$.

$$\begin{aligned}\dot{V}(z) &= 2x^T P(Ax + q(x)) \\ &= x^T (A^T P + PA)x + 2x^T Pq(x) \\ &\leq -x^T x + 2\alpha\|P\|\|x\|^2 \\ &= -(1 - 2\alpha\|P\|)\|x\|^2.\end{aligned}$$

If $1 - 2\alpha\|P\| > 0$, the system is globally asymptotically stable, since then we have

$$\dot{V}(z) \leq -(1 - 2\alpha\|P\|)\|x\|^2 \leq -\frac{1 - 2\alpha\|P\|}{\lambda_{\max}(P)} V(x).$$

Hence, we can take $\bar{\alpha} = 0.49/\|P\|$, since any $\alpha \leq \bar{\alpha}$ will satisfy $1 - 2\alpha\|P\| > 0$.

6. *Stability analysis of system with intermitent failures.* We consider the system $\dot{x} = A(t)x$, where $A(t) = A_{\text{nom}} \in \mathbf{R}^{n\times n}$ when system is working (*i.e.*, in *nominal mode*) and $A(t) = A_{\text{fail}} \in \mathbf{R}^{n\times n}$ when system is not working (*i.e.*, in *failure mode*). The nominal system is stable, *i.e.*, $\dot{z} = A_{\text{nom}} z$ is stable.

   We suppose that the system fails in various failure episodes, which are the intervals of time during which $A(t) = A_{\text{fail}}$. We assume that no failure episode can last longer than T_1 seconds. (A failure can, however, last shorter than T_1 seconds.) We also assume that once a failure has finished, no new failure occurs for at least T_2 seconds. In other words, T_2 gives the minimum time between two successive failure episodes, *i.e.*, a minimum time between failures. (Again, it is possible that the time between two successive failures exceeds T_2.) To simplify things, you can assume that the first failure does not occur until at least time T_2, *i.e.*, the system starts off at $t = 0$ in a working period, that is at least T_2 seconds long. We let $\alpha = T_1/(T_1 + T_2)$, which is an upper bound on the fraction of time the system can be in failure mode.

   Intuition suggests that if the failures occur for only a small fraction of time, then the system should still be globally asymptotically stable, *i.e.*, if α is small enough, the system is globally asymptotically stable. The goal in this problem is to verify and quantify this statement.

   Find a positive number $\bar{\alpha}$ such that for $\alpha \leq \bar{\alpha}$, you can guarantee global asymptotic stability of the system. The number $\bar{\alpha}$ must be explicit, *i.e.*, it should be easily calculated using standard matrix operations (such as computing eigenvalues and singular values, solving Lyapunov or Riccati equations, etc.), starting from the problem data A_{nom}, A_{fail}, T_1, T_2, and α.

Remark: to save you some trouble, we should point out a common misperception. Many people assume that the 'worst' sequence of failure events is to have the system fail for the maximum possible time, *i.e.*, T_1, then work normally for the smallest possible time, *i.e.*, T_2, and then repeat this pattern. *This assumption is false.*

Solution: We solve the Lyapunov equation $A_{\text{nom}}^T P + P A_{\text{nom}} + I = 0$, to get $P > 0$. We'll use Lyapunov function $V(z) = z^T P z$.

During the working period, $\dot{V}(x) = -||x||^2$. Let $\delta_1 = \frac{1}{\lambda_{\max}(P)} > 0$, we have $\dot{V} \leq -\delta_1 V$. This means that if the system is working over the interval $[t_1, t_2]$, we have

$$V(x(t_2)) \leq e^{-\delta_1(t_2 - t_1)} V(x(t_1)).$$

Since $\delta_1 > 0$, this gives a guranteed decrease in V.

Also, during the failure period, $\dot{V}(x) = x^T(A_{\text{fail}}^T P + P A_{\text{fail}})x$. Let $\delta_2 = \frac{\lambda_{\max}(A_{\text{fail}}^T P + P A_{\text{fail}})}{\lambda_{\min}(P)}$, we have $\dot{V} \leq \delta_2 V$. This means that if the system is in failure mode over the interval $[t_1, t_2]$, we have

$$V(x(t_2)) \leq e^{\delta_2(t_2 - t_1)} V(x(t_1)).$$

Since δ_2 can be positive, this allows for the possibility of V increasing over the period. But it gives a maximum possible increase in V.

Now let's put these together. At time t, we have

$$\begin{aligned} V(x(t)) &\leq V(x(0)) e^{-\delta_1 S_1} e^{\delta_2 S_2} \\ &= V(x(0)) e^{-t(\delta_1 \frac{S_1}{t} - \delta_2 \frac{S_2}{t})}, \end{aligned}$$

where S_1 is the total time the system has been working up to time t, and S_2 is the total time the system has failed up to time t.

Since $\frac{S_1}{t} \geq (1 - \alpha)$, $\frac{S_2}{t} \leq \alpha$, and $\delta_1 > 0$, $\delta_2 \leq \max(0, \delta_2)$, we have $\delta_1 \frac{S_1}{t} - \delta_2 \frac{S_2}{t} \geq \delta_1(1 - \alpha) - \max(0, \delta_2)\alpha$. Hence

$$V(x(t)) \leq V(x(0)) e^{-t(\delta_1(1-\alpha) - \max(0, \delta_2)\alpha)}$$

We see that V (and hence x) converges to zero provided

$$\delta_1(1 - \alpha) - \max(0, \delta_2)\alpha > 0,$$

i.e., provided $\alpha < \bar{\alpha}$, where

$$\bar{\alpha} = \frac{\delta_1}{\delta_1 + \max(0, \delta_2)}.$$

There were many incorrect proofs. Some of the incorrect proofs were based on the following idea. We look at the state whenever it enters a working mode (say). We can express the state as

$$x(t_{i+1}) = e^{A_{\text{fail}} a_i} e^{A_{\text{nom}} b_i} x(t_i),$$

where t_i is the ith time the system enters the nominal mode, and a_i and b_i are the times the system is failed and working, respectively, during the ith cycle. At this point, people made several errors. One was to state that if the magnitude of the eigenvalues of

$$e^{A_{\text{fail}} a_i} e^{A_{\text{nom}} b_i}$$

are less than one, the system is stable. This is false: it basically says that a time-varying system is stable if the "frozen" matrices $A(t)$ are all stable, which is false.

It was also incorrect to argue that

$$\left\| e^{A_{\text{nom}} b_i} \right\|$$

was less then one. This might not be the case, depending on b_i.