Stephen Boyd

Ask me about anything that isn't clear.

Linear dynamical system $\dot{x} = Ax$, with $x(t) \in \mathbf{R}^n$

System is called **constant norm** if for every trajectory x, $\|x(t)\|$ is constant, *i.e.*, doesn't depend on t

System is called **constant speed** if for every trajectory x, $\|\dot{x}(t)\|$ is constant, *i.e.*, doesn't depend on t

Give an example of a constant norm system.

Give an example of a constant speed system.

Find the conditions on A under which the system is constant norm.

Find the conditions on A under which the system is constant speed.

Is every constant norm system a constant speed system?

Is every constant speed system a constant norm system?

Discussion/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 $\dot{x} = (x_2(0), 0)$.

Stephen Boyd (1)

Ask me about anything that isn't clear.

$\mathbf{card}(x)$ denotes the number of nonzero entries in the vector $x \in \mathbf{R}^n$

suppose we have

$$y = Ax, \qquad \mathbf{card}(x) \leq k$$

you know $A \in \mathbf{R}^{m \times n}$, $y \in \mathbf{R}^m$, and k

how would you determine whether there is a **unique** x that satisfies these conditions, and if so, find x?

Discussion/solution.

Without the cardinality condition, there is a unique solution x only if A has zero nullspace. This requires that $m \geq n$ and that A have rank n. When we add the cardinality information, it can happen that we have a unique solution, even when $m < n$, or $\mathbf{Rank}(A) < n$. These ideas are central to a new research area called compressed sensing. But back to our problem ...

Let's consider *any* set of k indices. Form the matrix $\tilde{A} \in \mathbf{R}^{m \times k}$, taking only the associated columns of A. Now consider the equation $\tilde{A}z = y$. Any solution of this equation gives us a solution x of $Ax = y$, with $\mathbf{card}(x) \leq k$, just by inserting the entries of z into the positions of x associated with the indices, with zeros elsewhere. If the equation $\tilde{A}z = y$ has more than one solution, then the original x is not recoverable; there are at least two values of x that satisfy $Ax = y$ and $\mathbf{card}(x) \leq k$ (indeed, the two solutions have the same sparsity pattern). So the equation $\tilde{A}x = y$ can have only one or zero solutions. If $\tilde{A}x = y$ has one solution, then it is for sure a candidate for x.

Now, we carry out this analysis of the equation $\tilde{A}z = y$ for *all* $\binom{n}{k}$ choices of k indices from $1, \ldots, n$. If for any choice of indices there is more than one solution, we can't recover x. We can just quit the whole process right there.

If for all choices that have a solution, the solution is the same, then that vector is x, and it is the unique solution.

There are several ways to carry out this method. (There are also several incorrect ways to do it.) Here is one correct way: For each subset, check if $\tilde{A}z = y$ has a solution. If not, go on to the next subset. If it does, check the rank of $\tilde{A}$. If it is less than k, quit the entire algorithm, announcing

Now, this isn't really practical, since $\binom{n}{k}$ is a really big number, unless k is very small. But I didn't ask for a practical method.

None of the following was needed, but you might find it interesting. It is likely there isn't a much better way to answer the question with certainty than to do an exhaustive search over subset of cardinality k. However, there are some very good heuristics for finding a sparse x that satisfies $Ax = y$. One way is to minimize $\|x\|_1$ subject to $y = Ax$. This can be done using linear programming. This is a heuristic — it can be wrong — but it very often does recover a sparse x from $y = Ax$.

Ask me about anything that isn't clear.

The *average* of a vector $x \in \mathbf{R}^n$ is defined as

$$\mathbf{avg}(x) = \frac{x_1 + \cdots + x_n}{n}.$$

Average-preserving linear transformation.
Under what conditions on $A \in \mathbf{R}^{m \times n}$ do we have

$$\mathbf{avg}(Ax) = \mathbf{avg}(x)$$

for all $x \in \mathbf{R}^n$?

Average-reducing linear transformation.
Under what conditions on $A \in \mathbf{R}^{m \times n}$ do we have

$$|\mathbf{avg}(Ax)| \leq |\mathbf{avg}(x)|$$

for all $x \in \mathbf{R}^n$?

Discussion/solution. We can write $\mathbf{avg}(x) = (1/n)\mathbf{1}^T x$, so

$$\mathbf{avg}(Ax) = \mathbf{avg}(x) \iff (1/m)\mathbf{1}^T Ax = (1/n)\mathbf{1}^T x. \iff \mathbf{1}^T Ax = (m/n)\mathbf{1}^T x.$$

This holds for all x if and only if $\mathbf{1}^T A = (m/n)\mathbf{1}^T$, which can be expressed as $A^T\mathbf{1} = (m/n)\mathbf{1}$. This means that all columns of A must sum to m/n.

Another way to say it is: If you add up the rows of A, you get a row vector all of whose entries as m/n.

If A is square (which it need not be), the condition also means that A has $\mathbf{1}$ as a left eigenvector, with associated eigenvalue m/n.

The second question is a bit trickier. The solution is: $|\mathbf{avg}(Ax)| \leq |\mathbf{avg}(x)|$ for all x if and only if $A^T\mathbf{1} = \alpha(m/n)\mathbf{1}$ for some α with $|\alpha| \leq 1$. In other words, all columns of A must sum to m/n, times a constant (which is the same for all columns) less than or equal to one in magnitude. In terms of eigenvectors, the condition can be expressed as: A has $\mathbf{1}$ as a left eigenvector, with associated eigenvalue λ, with $|\lambda| \leq m/n$.

The "if" direction is clear: If $A^T\mathbf{1} = \alpha(m/n)\mathbf{1}$, where $|\alpha| \leq 1$, then for any x we have

$$\mathbf{avg}(Ax) = (1/m)|(A^T\mathbf{1})^T x| = (|\alpha|/n)|\mathbf{1}^T x| \leq (1/n)|\mathbf{1}^T x| = \mathbf{avg}(x).$$

Now we'll show the opposite direction. Let $a = A^T\mathbf{1}$ and $b = (m/n)\mathbf{1}$. Then $|\mathbf{avg}(Ax)| \leq |\mathbf{avg}(x)|$ can be written as $|a^T x| \leq |b^T x|$. Suppose that $|a^T x| \leq |b^T x|$ for all x. We'll show that $a = \alpha b$, for some $\alpha \in [-1, 1]$. Note that this holds if $a = 0$, with $\alpha = 0$, so we will assume that $a \neq 0$.

Clearly if $b^T x = 0$, then $a^T x = 0$. Thus $\mathcal{N}(b^T) \subseteq \mathcal{N}(a^T)$. Taking orthogonal complements we get $\mathcal{R}(b) \supseteq \mathcal{R}(a)$. In particular $b \in \mathcal{R}(a)$, which means that $b = \alpha a$ for some $\alpha \in \mathbf{R}$. Taking $x = b$ in $|a^T x| \leq |b^T x|$ yields

$$|a^T b| = |\alpha| a^T a \leq |b^T b| = \alpha^2 a^T a,$$

so $|\alpha| \leq \alpha^2$. From this we conclude $|\alpha| \leq 1$.

Another way to come to the conclusion that the sums of the columns of A must be equal is to consider the particular values of x given by $x = e_i - e_j$, with $i \neq j$. Then $\mathbf{avg}(x) = 0$, so we have to have $\mathbf{avg}(Ax) = 0$. But $\mathbf{avg}(Ax)$ is exactly half the difference of the sum of column i and column j. We conclude that these column sums must be equal; since i and j were arbitrary, we see that all columns of A must have the same sum.

The entries of an invertible $n \times n$ matrix A are integers.

When are all the entries of A^{-1} integers?

(Always? Never? Sometimes?)

Discussion/solution.

As always, the point is not the solution; the point is the clarity of the arguments used.

The identity matrix is an example showing it's possible for all entries of A and A^{-1} to be integers. Another more interesting example is an upper or lower triangular matrix, with its diagonal entries all 1 or -1.

Let's start with 1×1 matrices, $i.e.$, scalars. Here the inverse is an integer only if $A = 1$ or $A = -1$.

Now let's look at 2×2 matrices. We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

so if $\det A = 1$ or -1, then all entries of A^{-1} are integers. The converse is also true: if all entries of the inverse are integers, then $\det A = 1$ or -1. To see this, we note that

$$1 = \det I = \det(AA^{-1}) = (\det A)(\det A^{-1}).$$

If A and A^{-1} have all integer entries, then $\det A$ and $\det A^{-1}$ are both integers (since they are sums of products of entries). These two integers have a product equal to 1, so they can only be both 1, or both -1.

Now we can guess the general case: A^{-1} has integer entries if and only if $\det A$ is 1 or -1. To show one way, assume that $\det A = 1$ or -1. Cramer's formula for the inverse is

$$(A^{-1})_{ij} = \frac{(-1)^{i+j} \det \tilde{A}}{\det A},$$

where $\tilde{A}$ is formed from A by removing a column and a row. The numerator is an integer, and the denominator is 1 or -1, so $(A^{-1})_{ij}$ is an integer. To prove the converse, the argument above works: if A and A^{-1} both have integer entries, then $(\det A)(\det A^{-1})1$, and we conclude that $\det A = 1$ or -1.