# CS/ECE/ME 532

## Homework 1: Matrices, vectors, and norms

due: Friday September 16, 2016

**1. Matrix multiplication.** The local factory makes widgets and gizmos. Making one widget requires 3 lbs of materials, 4 parts, and 1 hour of labor. Making one gizmo requires 2 lbs of materials, 3 parts, and 2 hours of labor.

**a)** Write the information above in a matrix. What do the rows and columns represent?

**SOLUTION:** One possible representation is:

$$A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix}$$

The rows represent the requirements for widgets (first row) and gizmos (second row). Each is a row vector that contains the lbs of materials, number of parts, and hours of labor required. The columns represent the lbs of materials (first column), the number of parts (second column) and the hours of labor (third column). Each column is a vector where the first component corresponds to widgets and the second component corresponds to gizmos.

**b)** Suppose materials cost \$1/lb, parts cost \$10 each, and labor costs \$100/hr. Write this information in a vector. Write out a matrix-vector multiplication that calculates the total cost of making widgets and gizmos.

**SOLUTION:** Write the material costs in a column vector: $c = \begin{bmatrix} 1 & 10 & 100 \end{bmatrix}^\mathsf{T}$. The matrix multiplication that calculates total cost is:

$$t = Ac = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \\ 100 \end{bmatrix} = \begin{bmatrix} 143 \\ 232 \end{bmatrix}$$

So the total cost of making widgets and gizmos is \$143 and \$232 respectively.

**c)** Suppose the factory receives an order for 3 widgets and 4 gizmos. Again using matrix multiplication, find the total material, parts, and labor required to fill the order.

**SOLUTION:** Write the order in a column vector: $g = \begin{bmatrix} 3 & 4 \end{bmatrix}^\mathsf{T}$. The matrix multiplication that calculates the requirements is:

$$r = g^\mathsf{T} A = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 17 & 24 & 11 \end{bmatrix}$$

So to make 3 widgets and 4 gizmos, we would require 17 lbs of materials, 24 parts, and 11 hours of labor.

**d)** Calculate the total cost for the order (using, you guessed it, matrix multiplication)

**SOLUTION:** The total cost in dollars is $g^\mathsf{T} A c = 1357$

**2. Linear dynamical systems.** Linear dynamical systems are a popular way of modeling mechanical and electrical systems. In general, the model takes the form:

$$\begin{aligned} \boldsymbol{x}(t+1) &= A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= C\boldsymbol{x}(t) + D\boldsymbol{u}(t) \end{aligned} \qquad \text{for } t = 0, 1, \dots, N$$

For example, in an engine model, the inputs $\boldsymbol{u}(0), \boldsymbol{u}(1), \dots, \boldsymbol{u}(N) \in \mathbb{R}^m$ could represent the throttle, fuel, and air injected at each timestep and the outputs $\boldsymbol{y}(0), \boldsymbol{y}(1), \dots, \boldsymbol{y}(N) \in \mathbb{R}^k$, could represent the engine RPM and torque at each timestep. The matrices $A, B, C, D$ characterize the complicated dependence of the outputs on the inputs and $\boldsymbol{x}(0), \boldsymbol{x}(1), \dots, \boldsymbol{x}(N) \in \mathbb{R}^n$ are internal *state* variables. Note that $n$ might be quite large, even if $m$ and $k$ are small! Find a matrix $G$ that satisfies:

$$\begin{bmatrix} \boldsymbol{y}(0) \\ \vdots \\ \boldsymbol{y}(N) \end{bmatrix} = G \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{u}(0) \\ \vdots \\ \boldsymbol{u}(N) \end{bmatrix} \qquad \text{where } G \in \mathbb{R}^{k(N+1)\times(n+m(N+1))}$$

Note that $G$ should only depend on $A, B, C, D$; it should not contain any $\boldsymbol{u}$'s, $\boldsymbol{y}$'s, or $\boldsymbol{x}$'s.

**SOLUTION:** By substituting $\boldsymbol{x}(k)$ into the expression for $\boldsymbol{x}(k+1)$ iteratively, we may write out expressions for $\boldsymbol{x}(0), \dots, \boldsymbol{x}(N)$ as functions of just $\boldsymbol{x}(0)$ and $\boldsymbol{u}(0), \dots, \boldsymbol{u}(N-1)$:

$$\begin{aligned} \boldsymbol{x}(0) &= \boldsymbol{x}(0) \\ \boldsymbol{x}(1) &= A\boldsymbol{x}(0) + B\boldsymbol{u}(0) \\ \boldsymbol{x}(2) &= A^2\boldsymbol{x}(0) + AB\boldsymbol{u}(0) + B\boldsymbol{u}(1) \\ &\vdots \\ \boldsymbol{x}(N) &= A^N\boldsymbol{x}(0) + A^{N-1}B\boldsymbol{u}(0) + A^{N-2}B\boldsymbol{u}(1) + \dots + AB\boldsymbol{u}(N-2) + B\boldsymbol{u}(N-1) \end{aligned}$$

Now apply the definition of $\boldsymbol{y}(t)$ to each equation and obtain:

$$\begin{aligned} \boldsymbol{y}(0) &= C\boldsymbol{x}(0) + D\boldsymbol{u}(0) \\ \boldsymbol{y}(1) &= CA\boldsymbol{x}(0) + CB\boldsymbol{u}(0) + D\boldsymbol{u}(1) \\ \boldsymbol{y}(2) &= CA^2\boldsymbol{x}(0) + CAB\boldsymbol{u}(0) + CB\boldsymbol{u}(1) + D\boldsymbol{u}(2) \\ &\vdots \\ \boldsymbol{y}(N) &= CA^N\boldsymbol{x}(0) + CA^{N-1}B\boldsymbol{u}(0) + CA^{N-2}B\boldsymbol{u}(1) + \dots + CB\boldsymbol{u}(N-1) + D\boldsymbol{u}(N) \end{aligned}$$

Or, as a single matrix equation, this is:

$$\begin{bmatrix} \boldsymbol{y}(0) \\ \boldsymbol{y}(1) \\ \boldsymbol{y}(2) \\ \vdots \\ \boldsymbol{y}(N) \end{bmatrix} = \underbrace{\begin{bmatrix} C & D & 0 & 0 & \dots & 0 \\ CA & CB & D & 0 & \dots & 0 \\ CA^2 & CAB & CB & D & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ CA^N & CA^{N-1}B & \dots & CAB & CB & D \end{bmatrix}}_{G} \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{u}(0) \\ \boldsymbol{u}(1) \\ \vdots \\ \boldsymbol{u}(N) \end{bmatrix}$$

Several students asked whether there was a more elegant way to solve this problem rather than to expand every entry as above. There is! Read on for an alternate solution.

**Alternate approach** Note: you are **not** required to know this approach; it's just to satisfy your curiosity! We'll start by defining block versions of the matrices and vectors. Define:

$$x = \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{x}(1) \\ \vdots \\ \boldsymbol{x}(N) \end{bmatrix} \qquad u = \begin{bmatrix} \boldsymbol{u}(0) \\ \boldsymbol{u}(1) \\ \vdots \\ \boldsymbol{u}(N) \end{bmatrix} \qquad y = \begin{bmatrix} \boldsymbol{y}(0) \\ \boldsymbol{y}(1) \\ \vdots \\ \boldsymbol{y}(N) \end{bmatrix}$$

Also define the $(N+1) \times 1$ block vector and the $(N+1) \times (N+1)$ block matrices:

$$\bar{E} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \bar{S} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ I & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & I & 0 \end{bmatrix} \qquad \bar{A} = \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A \end{bmatrix} \qquad \text{and similarly for } \bar{B}, \bar{C}, \bar{D}$$

$\bar{S}$ is called the *shift* matrix because when you multiply it by a vector, you get a version of the vector that is shifted down, with a 0 added for the first entry and the last entry truncated. For example,

$$\bar{S}x = \begin{bmatrix} 0 \\ \boldsymbol{x}(0) \\ \vdots \\ \boldsymbol{x}(N-1) \end{bmatrix}$$

It's not too difficult to see (check for yourself!) that we can write our entire dynamics as the following compact set of equations:

$$\begin{aligned} x &= \bar{S}\left(\bar{A}x + \bar{B}u\right) + \bar{E}\boldsymbol{x}(0) \\ y &= \bar{C}x + \bar{D}u \end{aligned}$$

Note that the first equation contains $x$ in two places. We can solve for $x$ and then substitute into $y$. The result is:

$$y = \left(\bar{C}\left(I - \bar{S}\bar{A}\right)^{-1}\bar{B}\bar{S} + \bar{D}\right)u + \bar{C}\left(I - \bar{S}\bar{A}\right)^{-1}\bar{E}\boldsymbol{x}(0)$$

We set out to express $y$ in terms of $u$ and $\boldsymbol{x}(0)$, so this is it! We conclude that:

$$G = \left[\bar{C}\left(I - \bar{S}\bar{A}\right)^{-1}\bar{E} \quad \bar{C}\left(I - \bar{S}\bar{A}\right)^{-1}\bar{B}\bar{S} + \bar{D}\right]$$

How do we actually evaluate this? We can observe that:

$$I - \bar{S}\bar{A} = \begin{bmatrix} I & 0 & \dots & 0 \\ -A & I & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -A & I \end{bmatrix} \quad \text{and also:} \quad \left(I - \bar{S}\bar{A}\right)^{-1} = \begin{bmatrix} I & 0 & 0 & \dots & 0 \\ A & I & 0 & \dots & 0 \\ A^2 & A & I & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ A^N & \dots & A^2 & A & I \end{bmatrix}$$

Substituting this into our expression for $G$, we obtain the same result as before:

$$G = \begin{bmatrix} C & D & 0 & 0 & \dots & 0 \\ CA & CB & D & 0 & \dots & 0 \\ CA^2 & CAB & CB & D & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ CA^N & CA^{N-1}B & \dots & CAB & CB & D \end{bmatrix}$$

**3. Seminorm nonnegativity.** Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a function that satisfies the following properties:

- $f(a\boldsymbol{x}) = |a| f(\boldsymbol{x})$ for all $a \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$ (absolute homogeneity)
- $f(\boldsymbol{x}+\boldsymbol{y}) \le f(\boldsymbol{x}) + f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ (triangle inequality)

Use the properties above to prove that $f(\boldsymbol{x}) \ge 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

**SOLUTION:** By using the first property with $a = 0$ and $a = -1$, we conclude that:

$$f(\mathbf{0}) = 0 \qquad \text{and} \qquad f(-\boldsymbol{x}) = f(\boldsymbol{x})$$

By using the second property with $\boldsymbol{y} = \boldsymbol{x}$, we have:

$$\begin{aligned} & f(\boldsymbol{x}-\boldsymbol{x}) \le f(\boldsymbol{x}) + f(-\boldsymbol{x}) & \\ \implies \quad & f(\mathbf{0}) \le 2f(\boldsymbol{x}) & \text{(because } f(-\boldsymbol{x}) = f(\boldsymbol{x})\text{)} \\ \implies \quad & 0 \le f(\boldsymbol{x}) & \text{(because } f(\mathbf{0}) = 0\text{)} \end{aligned}$$

Note that this restricted set of functions $f$ (called seminorms) are a broader class than the set of norms. In other words, every norm is a seminorm but not every seminorm is a norm. It follows that since the nonnegativity result we proved here holds for all seminorms, it must also hold for all norms.

**4. Norm additivity.** Suppose $\|\cdot\|_a$ and $\|\cdot\|_b$ are norms on $\mathbb{R}^n$.

**a)** Prove that $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_a + \|\boldsymbol{x}\|_b$ is also a norm on $\mathbb{R}^n$.

**b)** Sketch the norm ball in $\mathbb{R}^2$ for the norm $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1 + \|\boldsymbol{x}\|_\infty$.

**SOLUTION:**

**a)** We will prove each property of the norm.

- First, $f(\mathbf{0}) = \|\mathbf{0}\|_a + \|\mathbf{0}\|_b = 0 + 0 = 0$. Next, suppose that $f(\mathbf{0}) = 0$. Then $\|\boldsymbol{x}\|_a + \|\boldsymbol{x}\|_b = 0$. Since every norm is nonnegative, it must be the case that $\|\boldsymbol{x}\|_a = \|\boldsymbol{x}\|_b = 0$. It therefore follows that $\boldsymbol{x} = 0$. In conclusion, $f(\boldsymbol{x}) = 0$ if and only if $\boldsymbol{x} = 0$.
- Absolute homogeneity: let $a \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$. Then, since every norm satisfies absolute homogeneity, we have: $f(a\boldsymbol{x}) = \|a\boldsymbol{x}\|_a + \|a\boldsymbol{x}\|_b = |a|\,\|\boldsymbol{x}\|_a + |a|\,\|\boldsymbol{x}\|_b = |a| f(\boldsymbol{x})$.
- Triangle inequality: apply the triangle inequality to each norm separately!

$$\begin{aligned} f(\boldsymbol{x}+\boldsymbol{y}) &= \|\boldsymbol{x}+\boldsymbol{y}\|_a + \|\boldsymbol{x}+\boldsymbol{y}\|_b \\ &\le (\|\boldsymbol{x}\|_a + \|\boldsymbol{y}\|_a) + (\|\boldsymbol{x}\|_b + \|\boldsymbol{y}\|_b) \\ &= (\|\boldsymbol{x}\|_a + \|\boldsymbol{x}\|_b) + (\|\boldsymbol{y}\|_a + \|\boldsymbol{y}\|_b) \\ &= f(\boldsymbol{x}) + f(\boldsymbol{y}) \end{aligned}$$

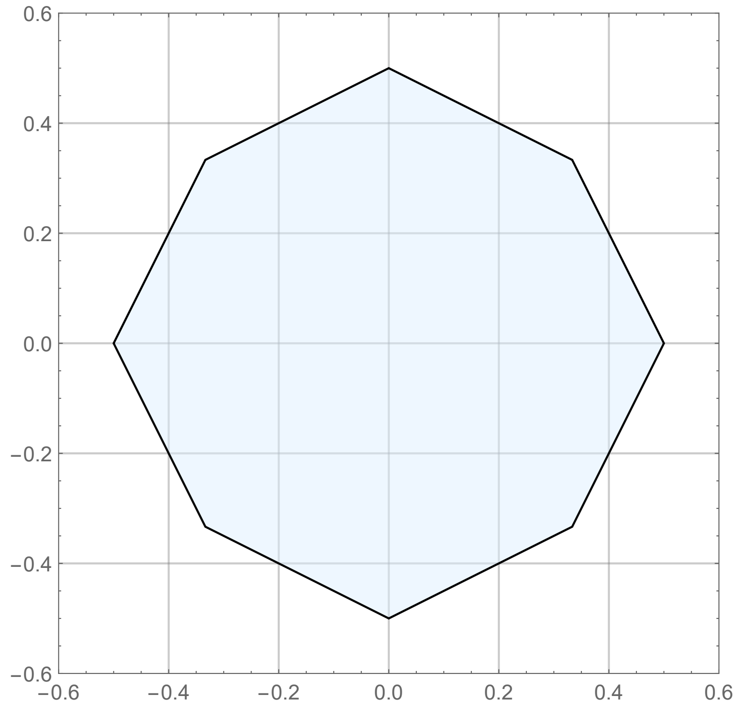
**b)** from part (a), we know that the sum of the 1-norm and $\infty$-norm must also be a norm. Let $\boldsymbol{x} = \begin{bmatrix} x & y \end{bmatrix}^\mathsf{T}$. By definition, we have:

$$f(\boldsymbol{x}) = |x| + |y| + \max(|x|, |y|)$$

Clearly $f$ has $x \mapsto -x$ symmetry as well as $y \mapsto -y$ symmetry. So it's enough to consider the first quadrant ($x \ge 0$ and $y \ge 0$) and then reflect the picture about the $x$ and $y$ axes to get the other three quadrants. So if we assume $x \geq 0$ and $y \geq 0$, there are two cases of interest. The result is that if $f(\boldsymbol{x}) \leq 1$, then

$$\begin{cases} 2x + y \leq 1 & \text{if } x \geq y \geq 0 \\ x + 2y \leq 1 & \text{if } y \geq x \geq 0 \end{cases}$$

Here is a picture of the full norm ball:




It is an 8-sided polygon with vertices at $(\pm\frac{1}{2}, 0)$, $(0, \pm\frac{1}{2})$, and $(\pm\frac{1}{3}, \pm\frac{1}{3})$.