Math 222 Spring 2016, Linear Algebra Review

We consider 2×2 systems, that is, two linear algebraic equations in two unknowns. The results can be generalized to $n \times n$ systems but these will not be needed in the context of this course, where we consider systems of just two coupled first order linear ODEs.

1. The general case. The general 2×2 system is

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{1}$$

where all quantities are real, $\{a_{ij}\}$ and $\{b_i\}$ are given and we want to solve for $\{x_i\}$ with i and $j = 1, 2$. It is often useful to write this in matrix form as

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{2}$$

or with vector notation as

$$\boldsymbol{Ax} = \boldsymbol{b}\,, \quad \text{where } \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \text{and } \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{3}$$

Here $\boldsymbol{A} = (a_{ij})$ is a 2×2 matrix, with 2 horizontal rows and two vertical columns, and a_{ij} is the element (or entry) in the i^{th} row and the j^{th} column. Both $\boldsymbol{x}$ and $\boldsymbol{b}$ are 2×1 column vectors. In general, an $m \times n$ matrix has m rows and n columns.

On the left hand side of (2) or (3), $\boldsymbol{Ax}$ is an example of matrix multiplication, which is defined as follows:

Definition. If $\boldsymbol{A} = (a_{ij})$ is a $m \times n$ matrix and $\boldsymbol{B} = (b_{ij})$ is a $n \times r$ matrix, their product $\boldsymbol{AB}$ is an $m \times r$ matrix with element in the i^{th} row and j^{th} column given by multiplying each element in the i^{th} row of $\boldsymbol{A}$ by the corresponding element in the j^{th} column of $\boldsymbol{B}$.

In other words, the i, j^{th} element of $\boldsymbol{AB}$ is the dot product of the i^{th} row of $\boldsymbol{A}$ with the j^{th} column of $\boldsymbol{B}$, both considered as vectors of dimension n. Equivalently, if $\boldsymbol{C} = \boldsymbol{AB}$ then

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}\,.$$

Example 1. If $\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 2 & 4 \\ -1 & 2 \end{pmatrix}$ then $\boldsymbol{AB} = \begin{pmatrix} -1 & 10 \\ 5 & 6 \end{pmatrix}$. Notice however that $\boldsymbol{BA} = \begin{pmatrix} 10 & 2 \\ 3 & -5 \end{pmatrix}$. It is true in general that, even for square matrices, the matrix product $\boldsymbol{AB} \neq \boldsymbol{BA}$, so that the matrix product 'does not commute'.

We have the 'identity matrix' which has entries 1 along the 'leading diagonal' and 0 elsewhere. So $\boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In this case only, for any matrix $\boldsymbol{A}$ we have $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$.

2. Solution of a linear algebraic system. The next two examples show the different types of solution structure we can find when solving $\boldsymbol{Ax} = \boldsymbol{b}$.

Example 2. Solve

$$\begin{aligned} x_1 + 3x_2 &= 5 \quad \text{(i)} \\ 2x_1 - x_2 &= 3 \quad \text{(ii)}. \end{aligned}$$

The quickest and simplest way is to eliminate one of x_1 or x_2 then find the other by 'back substitution'. For example, we can form (ii)$-2\times$(i) to eliminate x_1 and find $-7x_2 = -7$, so that $x_2 = 1$. Substitution back in (i) then gives $x_1 = 5 - 3x_2 = 2$.

Notice that geometrically the two equations represent two straight lines in the x_1, x_2-plane that intersect in the one point (2,1).

Example 3. Solve, where α and β are parameters,

$$\begin{aligned} x_1 + 2x_2 &= \alpha \quad \text{(i)} \\ 3x_1 + 6x_2 &= \beta \quad \text{(ii)}. \end{aligned}$$

We can form (ii) $-3\times$(i) to eliminate x_1, and we find that

$$0x_2 = \beta - 3\alpha \quad \text{(iii)}.$$

This leads to two cases:

(a) If $\beta = 3\alpha$ then (iii) is satisfied for any x_2. We have $x_1 + 2x_2 = \alpha$ in (i), and equation (ii) is just 3×(i), which gives no further information. In the x_1, x_2-plane we have a line of solutions, or infinitely many solutions, each one being a point on the line. We can let $x_2 = t$ to find parametric equations for the line, as $x_1 = \alpha - 2t$ and $x_2 = t$. We have a solution for all t no matter what value α takes – when $\beta = 3\alpha$. Geometrically, (i) and (ii) are two *coincident* lines.

(b) If $\beta \neq 3\alpha$ then we can not satisfy (iii) for any x_2. In this case there is no solution for x_1 and x_2. Equations (i) and (ii) are said to be 'inconsistent'. Geometrically, (i) and (ii) are two parallel lines that are *not* coincident, and therefore do not intersect .

Homogeneous and nonhomogeneous. A system of linear algebraic equations $\boldsymbol{Ax} = \boldsymbol{b}$ is 'nonhomogeneous' when the right hand side $\boldsymbol{b} \neq \boldsymbol{0}$. When $\boldsymbol{b} = \boldsymbol{0}$ we have $\boldsymbol{Ax} = \boldsymbol{0}$ and the system is 'homogeneous'. Notice that this terminology is very similar to the use of the same words in the context of ODE's for $Lu = f$, where L is a differential operator and u and f are functions of t.

If we set $\boldsymbol{b} = \boldsymbol{0}$ in the last two examples, then in example 2 we find that there is just one solution again, and it is $\boldsymbol{x} = (x_1, x_2)^T = \boldsymbol{0}$. However, in example 3, we are in case (a), not (b), since $\beta = 3\alpha = 0$, and we have the nonzero solution $\boldsymbol{x} = t(-2, 1)^T$ for any t.

3. Return to the general case. In

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \quad \text{(i)} \\ a_{21}x_1 + a_{22}x_2 &= b_2 \quad \text{(ii)} \end{aligned} \tag{4}$$

we can eliminate x_1 by taking $a_{11}\times$(ii) $-a_{21}\times$(i) to find

$$(a_{11}a_{22} - a_{12}a_{21})\, x_2 = a_{11}b_2 - a_{21}b_1\,. \tag{5}$$

Back substitution with this works, but is slow, so we eliminate x_2 by taking $a_{22}\times$(i) $-a_{12}\times$(ii) to find

$$(a_{11}a_{22} - a_{12}a_{21})\, x_1 = a_{22}b_1 - a_{12}b_2\,. \tag{6}$$

On the left hand side of both expressions, we have the determinant of $\boldsymbol{A}$,

$$det\boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,,$$

and we can write the solution for x_1 and x_2 in vector form as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{det\boldsymbol{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{7}$$

You may recognize this as 'Cramer's rule' for solving linear algebraic equations. It can be generalized to larger $n\times n$ systems. It is OK to use it when $n = 2$, although it has no advantage over elimination and back substitution, and it becomes increasingly slow to compute, either by hand or numerically, as n increases. Even when $n = 3$ the method of elimination and back substitution is quicker. It is then referred to as Gaussian elimination, and the construction of linear combinations to eliminate components of $\boldsymbol{x}$ before back substitution is done by 'elementary row operations'.

Inverse of a matrix. We have constructed (7) because it makes an important point that is true for general $n \times n$ systems: Notice that we can only reason from (5) and (6) to (7) when the matrix $\boldsymbol{A}$ is such that $det\boldsymbol{A} \neq 0$. Also, when $det\boldsymbol{A} \neq \boldsymbol{0}$, we can write (7) in vector notation as

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}\,, \quad \text{where } \boldsymbol{A}^{-1} = \frac{1}{det\boldsymbol{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}, \tag{8}$$

where $\boldsymbol{A}^{-1}$ is called the 'inverse' of the matrix $\boldsymbol{A}$. As an exercise, we can check by matrix multiplication in the 2×2 case, that

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}\,. \tag{9}$$

4. Central Result. From (7) we can conclude that for the 2×2 case, and it is true in general when $\boldsymbol{A}$ is an $n \times n$ matrix and $\boldsymbol{x}$ and $\boldsymbol{b}$ are n-dimensional column vectors, that:

(i) When the matrix $\boldsymbol{A}$ is such that $det\boldsymbol{A} \neq 0$, the inverse matrix $\boldsymbol{A}^{-1}$ exists, and the solution of the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$. Both $\boldsymbol{A}^{-1}$ and $\boldsymbol{x}$ are unique. The matrix $\boldsymbol{A}$ is said to be 'nonsingular' or 'invertible'.
(ii) Conversely, when $\boldsymbol{A}$ is such that $det\boldsymbol{A} = 0$, there is no inverse matrix $\boldsymbol{A}^{-1}$. The matrix $\boldsymbol{A}$ is said to be 'singular' or 'non-invertible'.

In what follows below, we need the second part (ii). It is true in general that it is only when $det\boldsymbol{A} = 0$ that the *homogeneous* system $\boldsymbol{Ax} = \boldsymbol{0}$ has *nonzero* solutions. In the first part (i), with $\boldsymbol{b} = \boldsymbol{0}$ the only solution is $\boldsymbol{x} = \boldsymbol{0}$.

Example 4. We saw a contrast or difference in the form that the solution of a 2×2 system can take in examples (2) and (3) above. The difference is due to the value of the determinant $det\boldsymbol{A}$, and we have examples of each of the two cases (i) and (ii) of the central result (4) above.

In example (2), check that $det\boldsymbol{A} = 1(-1) - 3(2) = -7 \neq 0$, so we are in case (i). $\boldsymbol{A}$ is invertible, and we can find the inverse matrix from (8). It is

$$\boldsymbol{A}^{-1} = \begin{pmatrix} 1/7 & 3/7 \\ 2/7 & -1/7 \end{pmatrix}.$$

We can then compute $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ to find the unique solution or single point $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = (2, 1)^T$, but the elimination we performed earlier is probably quicker.

Conversely, in example (3), $det\boldsymbol{A} = 0$. We saw that the *non*homogeneous system $\boldsymbol{Ax} = \boldsymbol{b}$ only has a solution for a specific choice of $\boldsymbol{b}$, and then it is nonunique and is a line of points. For other $\boldsymbol{b} \neq \boldsymbol{0}$, the nonhomogeneous has system no solution. However, the *homogeneous* system $\boldsymbol{Ax} = \boldsymbol{0}$ has nonzero solutions, as we saw when we introduced the terms homogeneous and nonhomogeneous for algebraic equations.

5. Eigenvalues and eigenvectors of a 2 × 2 matrix.

We have seen that $\boldsymbol{Ax}$ is a vector, and when $\boldsymbol{A}$ is given or fixed it takes or maps a general point $\boldsymbol{x}$ in the plane to another point in the plane. We can think of the two points as the tips of two vectors, both with their tails at the origin. In this way a matrix can be thought of as a linear map or transformation of the plane that maps or transforms the vector $\boldsymbol{x}$ to the vector $\boldsymbol{Ax}$.

Example 5. The matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$ maps the unit vector $\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ that is parallel to the x-axis to $\boldsymbol{Ai} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and it maps the unit vector $\boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ that is parallel to the y-axis to $\boldsymbol{Aj} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$.

There are vectors in particular directions that remain unchanged by the action of $\boldsymbol{A}$, and they turn out to be important in many applications. For these, the vector $\boldsymbol{Ax}$ is parallel to $\boldsymbol{x}$, and so one is a scalar multiple of the other. We then have

$$\boldsymbol{Ax} = \lambda\boldsymbol{x} \quad \text{for} \quad \boldsymbol{x} \neq \boldsymbol{0} \text{ and some } \lambda\,. \tag{10}$$

Now, (10) is a linear algebraic system of equations. Its solutions for $\boldsymbol{x}$ are the called the 'eigenvectors' of $\boldsymbol{A}$ and the corresponding values of λ are the 'eigenvalues' of $\boldsymbol{A}$.

To find these, notice that for any vector $\boldsymbol{x}$ if we take the identity matrix $\boldsymbol{I}$ then $\boldsymbol{Ix} = \boldsymbol{x}$, since

$$\boldsymbol{Ix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \boldsymbol{x}.$$

So, we can write (10) as $\boldsymbol{Ax} = \lambda \boldsymbol{Ix}$, Next, put λ on the left hand side to find $\boldsymbol{Ax} - \lambda \boldsymbol{Ix} = \boldsymbol{0}$, and then group the terms in $\boldsymbol{x}$ to find $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$. Notice that, like $\boldsymbol{A}$, $\lambda \boldsymbol{I}$ is also a square 2×2 matrix. We now have the homogeneous algebraic system

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \quad \text{or} \quad \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{11}$$

It is a part of the definition of an eigenvector $\boldsymbol{x}$ at (10) that $\boldsymbol{x} \neq \boldsymbol{0}$, so we are looking for solutions of (11) that are nonzero. From our central result on linear algebra above, we therefore require that

$$det(\boldsymbol{A} - \lambda \boldsymbol{I}) = det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = 0. \tag{12}$$

This gives a quadratic equation for the eigenvalues λ that has real coefficients when the matrix $\boldsymbol{A}$ is real. It is

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0 \quad \text{or} \quad \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0. \tag{13}$$

This is called the 'characteristic polynomial' of $\boldsymbol{A}$, and for real $\boldsymbol{A}$ its two roots are either real or a complex conjugate pair. Once we have found the two eigenvalues λ_1 and λ_2, we return to the linear system (11) to find the corresponding eigenvectors $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$. We can also write (13) as

$$\lambda^2 - \lambda\, tr(\boldsymbol{A}) + det(\boldsymbol{A}) = 0\,,$$
$$\text{where } tr(\boldsymbol{A}) = a_{11} + a_{22} = \text{sum of diagonal entries of } \boldsymbol{A}\,,$$

which is called the 'trace of $\boldsymbol{A}$'.

Example 6. (Question 20 from the problems in section 7.3.) Find the eigenvalues and eigenvectors of the matrix $\boldsymbol{A} = \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$.
These are the solutions of $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$, and so

$$\begin{pmatrix} 1 - \lambda & \sqrt{3} \\ \sqrt{3} & -1 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

For solutions $\boldsymbol{x} \neq \boldsymbol{0}$ the characteristic polynomial is $det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$, so that the eigenvalues are given by

$$-(1 - \lambda)(1 + \lambda) - 3 = 0 \quad \Rightarrow \lambda^2 = 4 \ \Rightarrow \lambda_1 = 2, \text{ and } \lambda_2 = -2.$$

(a) With eigenvalue $\lambda_1 = 2$, $(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rl} -x_1 + \sqrt{3}x_2 = 0 & \text{(i)} \\ \sqrt{3}x_1 - 3x_2 = 0 & \text{(ii).} \end{array}$$

Equation (i) implies that $x_1 = \sqrt{3}x_2$, and equation (ii)$= -\sqrt{3}\times$(i). Let $x_2 = c_1$ be a parameter, then the eigenvector with eigenvalue $\lambda_1 = 2$ is $\boldsymbol{x}^{(1)} = c_1 \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}$. We can check that $\boldsymbol{A}\boldsymbol{x}^{(1)} = 2\boldsymbol{x}^{(1)}$.

(b) With eigenvalue $\lambda_2 = -2$, $(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rl} 3x_1 + \sqrt{3}x_2 = 0 & \text{(i)} \\ \sqrt{3}x_1 + x_2 = 0 & \text{(ii)}. \end{array}$$

Equation (i) implies that $x_1 = -x_2/\sqrt{3}$, and equation (ii)=(i)$/\sqrt{3}$. Let $x_1 = c_2$ be a parameter, then the eigenvector with eigenvalue $\lambda_2 = -2$ is $\boldsymbol{x}^{(2)} = c_2 \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}$. We can check that $\boldsymbol{A}\boldsymbol{x}^{(2)} = -2\boldsymbol{x}^{(2)}$

Example 7. In this example, the matrix $\boldsymbol{A}$ is real but the eigenvalues and eigenvectors are complex. Find the eigenvalues and eigenvectors of the matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix}$.

These are the solutions of $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$, and so

$$\begin{pmatrix} 1-\lambda & 1 \\ -2 & 3-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The characteristic polynomial is $det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$, so that the eigenvalues are given by

$$(1-\lambda)(3-\lambda) + 2 = 0 \quad \Rightarrow \lambda^2 - 4\lambda + 5 = 0 \ \Rightarrow \lambda_1 = 2 + i, \text{ and } \lambda_2 = 2 - i.$$

Since $\boldsymbol{A}$ is real the eigenvalues are a complex conjugate pair, i.e., $\lambda_2 = \overline{\lambda_1}$, where the bar denotes the complex conjugate $\overline{u + iv} = u - iv$.

(a) With eigenvalue $\lambda_1 = 2 + i$, $(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} -1-i & 1 \\ -2 & 1-i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{rl} -(1+i)x_1 + x_2 = 0 & \text{(i)} \\ -2x_1 + (1-i)x_2 = 0 & \text{(ii)}. \end{array}$$

Equation (i) implies that $x_2 = (1+i)x_1$, and equation (ii)$= (1-i)\times$(i). Let $x_1 = c_1$ be a parameter, then the eigenvector with eigenvalue $\lambda_1 = 2 + i$ is $\boldsymbol{x}^{(1)} = c_1 \begin{pmatrix} 1 \\ 1+i \end{pmatrix}$.

(b) With eigenvalue $\lambda_2 = \overline{\lambda_1} = 2 - i$, we can compute the eigenvector using the method above. Or, since λ_1 and $\boldsymbol{x}^{(1)}$ are an eigenvalue and eigenvector pair, $(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{x}^{(1)} = \boldsymbol{0}$. But $\boldsymbol{A}$ and $\boldsymbol{I}$ are real, so that $\boldsymbol{A} = \overline{\boldsymbol{A}}$, $\boldsymbol{I} = \overline{\boldsymbol{I}}$, and $\lambda_2 = \overline{\lambda_1}$. So, taking the complex conjugate of the relation for λ_1 and $\boldsymbol{x}^{(1)}$, we have $(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\overline{\boldsymbol{x}^{(1)}} = \boldsymbol{0}$, and so $\boldsymbol{x}^{(2)} = \overline{\boldsymbol{x}^{(1)}}$. The eigenvector with eigenvalue $\lambda_2 = 2 - i$ is therefore $\boldsymbol{x}^{(2)} = c_2 \begin{pmatrix} 1 \\ 1-i \end{pmatrix}$, where c_2 is a parameter.