THE MATRIX EQUATION $AX = B$ [1.4]

The product Ax

Definition: If A is an $m \times n$ matrix, with columns $a_1, ..., a_n$, and if x is in $\mathbb{R}^n$, then the product of A and x, denoted by Ax is the linear combination of the columns of A using the corresponding entries in x as weights; that is,

$$Ax = [a_1, a_2, \cdots, a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 a_1 + x_2 a_2 + \cdots x_n a_n$$

➤ Ax is defined only if the number of columns of A equals the number of entries in x

Example:

Let $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & -2 \\ 0 & -2 & 3 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ Then:

$$\boldsymbol{Ax} = 2 \times \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + 3 \times \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \\ 11 \end{bmatrix}$$

➤ $\boldsymbol{Ax}$ is the Matrix-by-vector product of $\boldsymbol{A}$ by $\boldsymbol{x}$

➤ 'matvec'

✍ What is the cost (operation count) of a 'matvec'?

Properties of the matrix-vector product

Theorem: If A is an $m \times n$ matrix, u and v are vectors in $\mathbb{R}^n$, and α is a scalar, then

1. $A(u + v) = Au + Av$;
2. $A(\alpha u) = \alpha(Au)$

✍ Prove this result using only the definition (columns)

✍ Prove that for any vectors u, v in $\mathbb{R}^n$ and any scalars α, β we have

$$A(\alpha u + \beta v) = \alpha Au + \beta Av$$

Row-wise matrix-vector product

➤ (in the form of an exercise)

➤ Suppose you have an $m \times n$ matrix A and a vector x of size n, show how you can compute an entry of the result $y = Ax$, without computing the others. Use the following example.

Example:

Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & -2 \\ 0 & -2 & 3 \end{bmatrix}$ and $x = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$. Let $y = Ax$

✍ How would you compute y_2 (only)

✍ Cost?

✍ General rule or process?

✍ Matlab code?

The matrix equation $Ax = b$

➤ We can now write a system of linear equations as a vector equation involving a linear combination of vectors.

➤ For example, the system $\boxed{\begin{aligned} x_1 + 2x_2 \quad -x_3 &= 4 \\ -\, 5x_2 + 3x_3 &= 1 \end{aligned}}$ is equivalent to

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

➤ The linear combination on the left-hand side is a matrix-vector product Ax with:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

➤ So: Can write above system as $Ax = b$ with $b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

➤ $\boldsymbol{Ax = b}$ is called a matrix equation.

! Used in textbook. Better terminology: "Linear system in matrix form"

➤ $\boldsymbol{A}$ is the coefficient matrix, $\boldsymbol{b}$ is the right-hand side

➤ So we have 3 different ways of writing a linear system

1. As a set of equations involving $\boldsymbol{x_1, ..., x_n}$
2. In an augmented matrix form
3. In the form of the matrix equation $\boldsymbol{Ax = b}$

➤ Important: these are just 3 different ways to look at the same equations. Nothing new. Only the notation is different.

Existence of a solution

➤ The equation $Ax = b$ has a solution if and only if b can be written as a linear combination of the columns of A

Theorem: Let A be an $m \times n$ matrix. Then the following four statements are all mathematically equivalent.

1. For each b in $\mathbb{R}^m$, the equation $Ax = b$ has a solution.
2. Each b in $\mathbb{R}^m$ is a linear combination of the columns of A.
3. The columns of A span $\mathbb{R}^m$
4. A has a pivot position in every row.

Proof

First: 1, 2, 3 are mathematically equivalent. They just restate the same fact which is represented by statement 2.

➤ So, it suffices to show (for an arbitrary matrix $\boldsymbol{A}$) that (1) is true iff (4) is true, i.e., that (1) and (4) are either both true or false.

➤ Given $\boldsymbol{b}$ in $\mathbb{R}^{m}$, we can row reduce the augmented matrix $[\boldsymbol{A}|\boldsymbol{b}]$ to reduced row echelon form $[\boldsymbol{U}|\boldsymbol{d}]$.

➤ Note that $\boldsymbol{U}$ is the rref of $\boldsymbol{A}$.

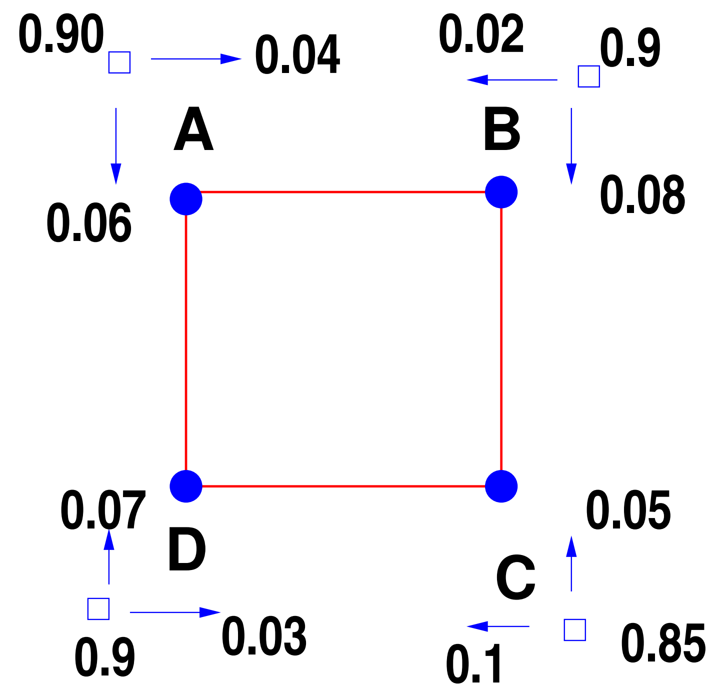
➤ If statement (4) is true, then each row of $\boldsymbol{U}$ contains a pivot position, and so $\boldsymbol{d}$ cannot be a pivot column.

➤ So $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution for any $\boldsymbol{b}$, and (1) is true.

- If (4) is false, then the last row of $\boldsymbol{U}$ is all zeros.
- Let $\boldsymbol{d}$ be any vector with a 1 in its last entry. Then $[\boldsymbol{U}|\boldsymbol{d}]$ represents an inconsistent system.
- Since row operations are reversible, $[\boldsymbol{U}|\boldsymbol{d}]$ can be transformed back into the form $[\boldsymbol{A}|\boldsymbol{b}]$.
- The new system $\boldsymbol{Ax} = \boldsymbol{b}$ is also inconsistent, and (1) is false. $\blacksquare$

Application: Markov Chains

Example: The annual population movement between four cities with an initial population of 1M each, follows the pattern shown in the figure: each number shows the fraction of the current population of city X moving to city Y. Migrations $A \leftrightarrow C$ and $B \leftrightarrow D$ are negligible.



- Is there an equilibrium reached?

- If so what will be the population of each city after a very long time?

➤ Let $\boldsymbol{x}^{(t)}$ = population distribution among cities at year $\boldsymbol{t}$ [starting at $\boldsymbol{t = 0}$] - no pop. growth is assumed.

✍ Express one step of the process as a matrix-vector product:

$$\boldsymbol{x}^{(t+1)} = \boldsymbol{A}\boldsymbol{x}^{(t)}$$

$$\boldsymbol{x}^{(t)} = \begin{bmatrix} \boldsymbol{x}_A^{(t)} \\ \boldsymbol{x}_B^{(t)} \\ \boldsymbol{x}_C^{(t)} \\ \boldsymbol{x}_D^{(t)} \end{bmatrix}$$

What is $\boldsymbol{A}$? What distinct properties does it have?

✍ Do one step of the process by hand.

✍ "Iterate" a few steps with matlab (40-50 steps)

✍ At the limit $\boldsymbol{Ax} = \boldsymbol{x}$, so $\boldsymbol{x}$ is the solution of a 'homogeneous' linear system. Find all possible solutions of this system. Among these which one is relevant?

✍ Compare with the solution obtained by "iteration"

Application: Leontief Model [sec. 1.6 of text]

- Equilibrium model of the economy
- Suppose we have 3 industries only [reality: hundreds]:

coal *electric* *steel*

- Each sector consumes output from the other two (+itself) and produces output that is in turn consumed by the others.

Distribution of Output from:			
Coal	Elec.	Steel	Purchased by
.0	.4	.6	Coal
.6	.1	.2	Elec.
.4	.5	.2	Steel
1	1	1	Total

- Problem: Find production quantities (called prices in text) of each of the 3 goods so that each sector's income matches its expenditure

- Expense for Coal: $.4p_E + .6p_S$ so we must have

$$\boxed{p_C = .4p_E + .6p_S \rightarrow p_C - .4p_E - .6p_S = 0}$$

- Similar reasoning for the other 2.
- In the end: Linear system of equations that is 'homogeneous' (RHS is zero).

$$\left[\begin{array}{rrr|r} 1 & -.4 & -.6 & 0 \\ -.6 & .9 & -.2 & 0 \\ -.4 & -.5 & .8 & 0 \end{array}\right]$$

✍ Use matlab to find general solution [Hint: Find the `rref` form first]