Motivation, Meaning and Context in Teaching Linear Algebra

David Strong
Pepperdine University

Examples to Motivate

Examples to motivate the entire course

- Good first day examples to motivate the entire course:
  - Interesting
  - Real
  - Simple
  - Involve (and thus introduce) several concepts
- Two uses of matrices:
  - Array of coefficients
  - Operator

A matrix as an array of coefficients: A nickel and dime problem

- A first day example. Choose a certain number of nickels and dimes:
    - Six coins
    - Five times as many dimes as nickels
    - 75 cents

A matrix as an array of coefficients:
A nickel and dime problem

- A first day example. Choose a certain number of nickels and dimes:
  - Six coins
  - Five times as many dimes as nickels
  - 75 cents

Problem	Condition(s)	Values of (n, d)
1	1	…, (0,6), (1,5), …, (6,0), …
2	2	…, (0,0), (1,5), …, (2,10), …
3	3	…, (1,7), (3,6), …, (15,0), …
4	1, 2	(1,5)
5	1, 3	(-3, 9)
6	2, 3	(15/11, 75/11)
7	1, 2, 3	-

The case for mathematics

- More students are now more curious.
- They also have begun to see the need for mathematics: language, equations, how to solve a system of equations, etc.

The case for mathematics

- More students are now more curious.
- They also have begun to see the need for mathematics: language, equations, how to solve a system of equations, etc.
- There are a plethora of issues, observations and questions that arise in this simple problem.

Issues: How to solve system of equations

$$\begin{aligned} n + \quad d &= 6 \\ 5n - \quad d &= 0 \\ 5n + 10d &= 75 \end{aligned}$$

- Students have some experience and ideas when # equations = # unknowns. Let's build on what they already know.
- There are multiple ways to solve the system of equations. Is there a preferred/best way?
  - What if there is a single solution? (What they are used to up to now.)
  - What if there are multiple solutions?
  - What if there is no solution?

Issues: Geometric/visualization

- Visualization is useful—we could plot the lines corresponding to the equations.
- But what if there are more than two unknowns?
  With three unknowns it can be a bit clumsy and less obvious.
  With four or more unknowns we simply cannot plot anything.

Issues: # unknowns vs. # equations

- While not completely obvious in just this example, students have their first experience with the very important idea in general:
  - # equations > # unknowns ⇒ no solution
  - # equations = # unknowns ⇒ one solution
  - # equations < # unknowns ⇒ infinite solutions

  And if there is more than one solution, there is an infinite number.
- Can there be exceptions to the above "rule"?
  - If # equations > # unknowns , can there still be a solution?
  - Can we end up with no solution if # equations = # unknowns ?
  - Etc.

Other issues, observations, questions

- Pivots, pivot rows, pivot columns, etc.
- Linear combinations, span, column space, linear independence:
  - In $A\vec{x} = \vec{b}$, can we build $\vec{b}$ out of the columns of A.
    That is, is $\vec{b}$ in the column space of A.
  - Do the columns of A span all of R^3?
  - What conditions are necessary for the columns of A to span R^3?
  - What conditions are sufficient for the columns of A to span R^3?
  - If $\vec{b}$ can be built out of the columns of A, is there more than one way to do so?
    If so, is this a “good” thing or a “bad” thing?

Other issues, observations, questions

- A "best" solution, if there is not an exact solution:
  - What does a "best" solution even mean?
  - Building a vector that is as close to the desired vector $\vec{b}$ using the given vectors, that is, the columns of A.
  - Projecting a vector onto the column space of a matrix or onto some other collection of vectors.
- Changing right hand side values possibly change existence and/or the number of solutions and the values of those solutions.
- We don't get into all of these issues right away, but all of these arise in this simple first day example, which we occasionally revisit throughout the course. A good first day example is one that is real, simple, interesting, and can be revisited later while/after learning additional concepts.

A matrix as an operator:
A (slightly contrived) discrete dynamical system

$$W_{k+1} = \frac{\sqrt{3}}{2} W_k + \frac{1}{2} R_k$$

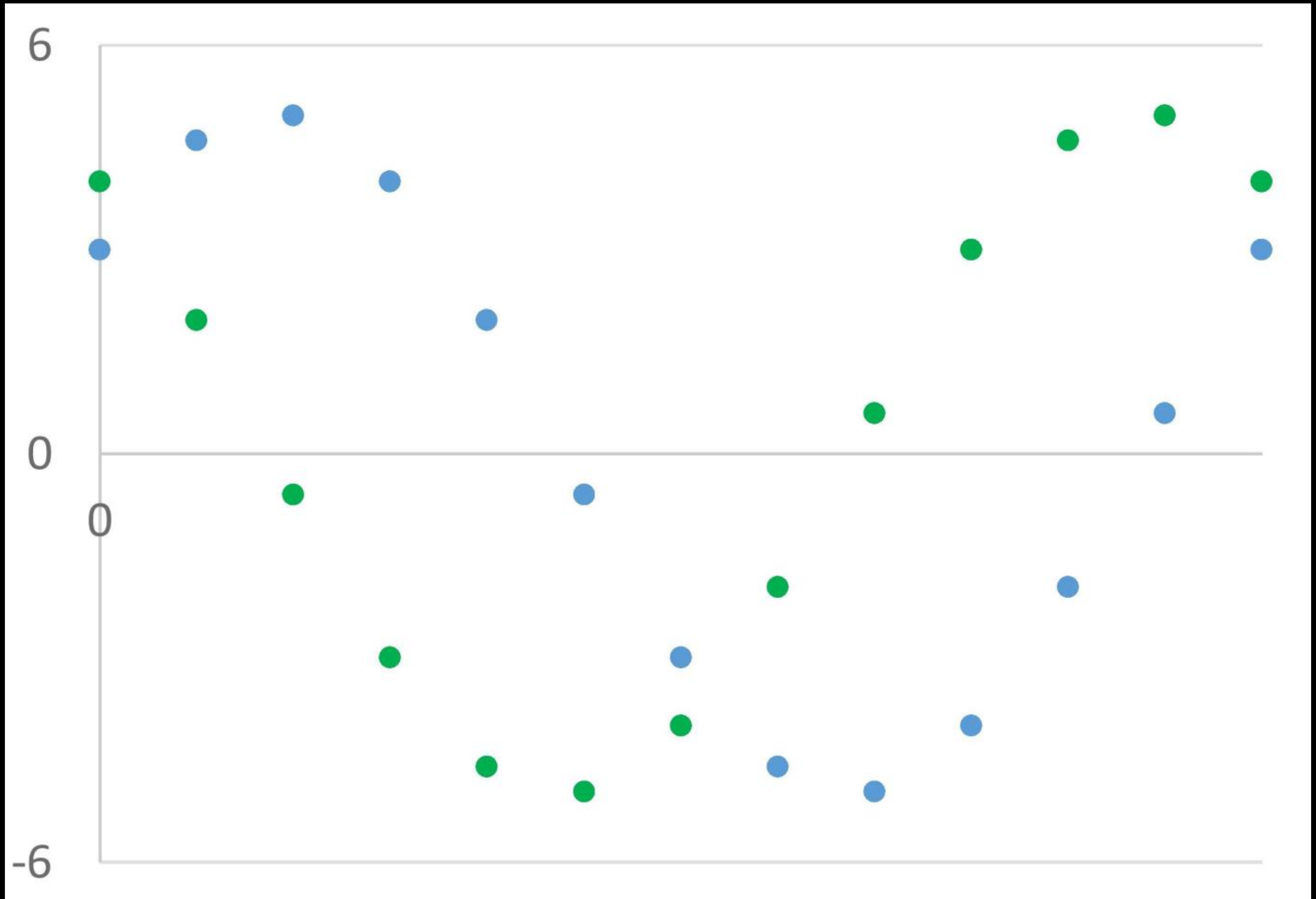
$$R_{k+1} = -\frac{1}{2} W_k + \frac{\sqrt{3}}{2} R_k$$

i.e.

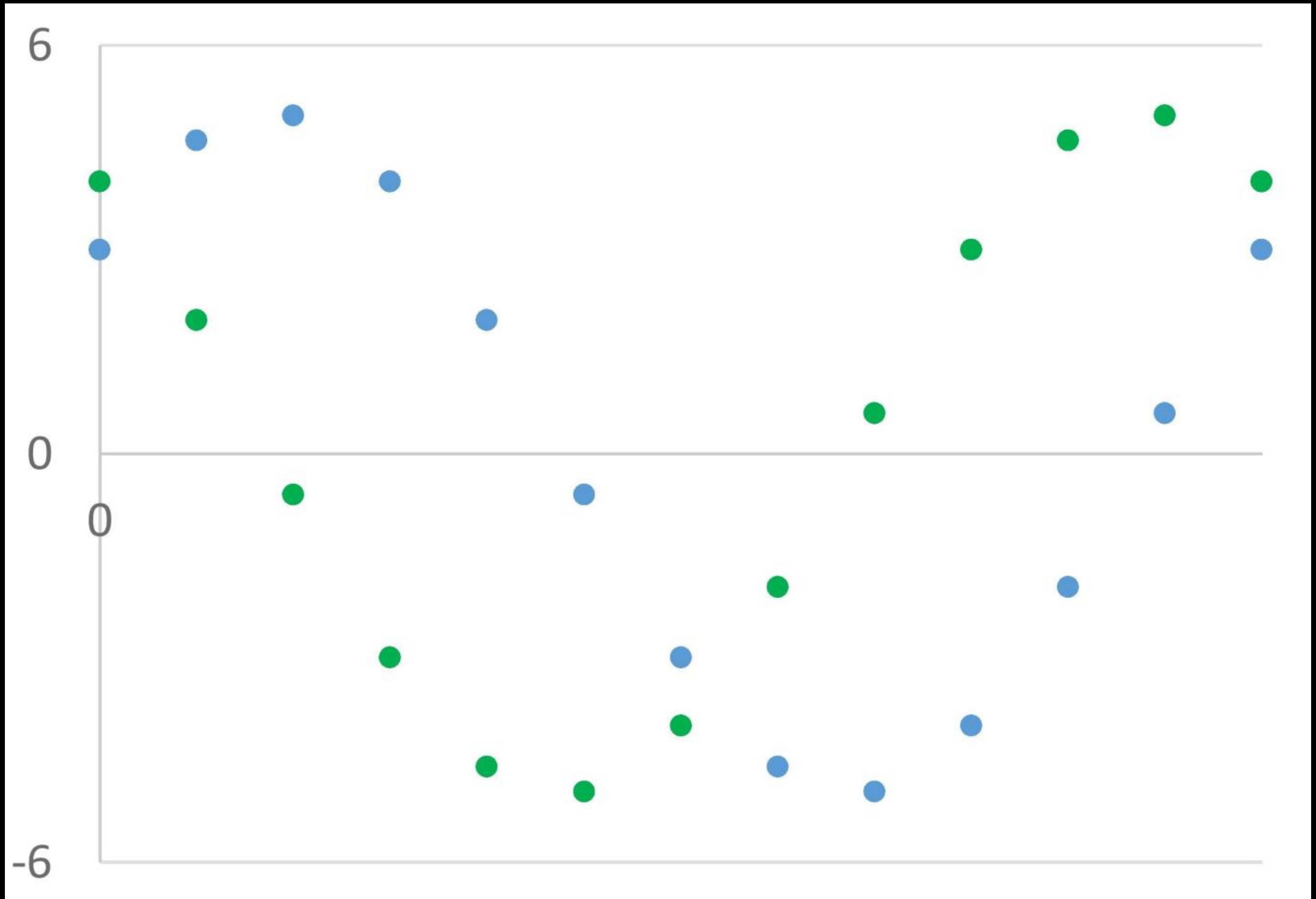
$$\begin{bmatrix} W \\ R \end{bmatrix}_{k+1} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} W \\ R \end{bmatrix}_k \approx \begin{bmatrix} .866 & .500 \\ -.500 & .866 \end{bmatrix} \begin{bmatrix} W \\ R \end{bmatrix}_k$$

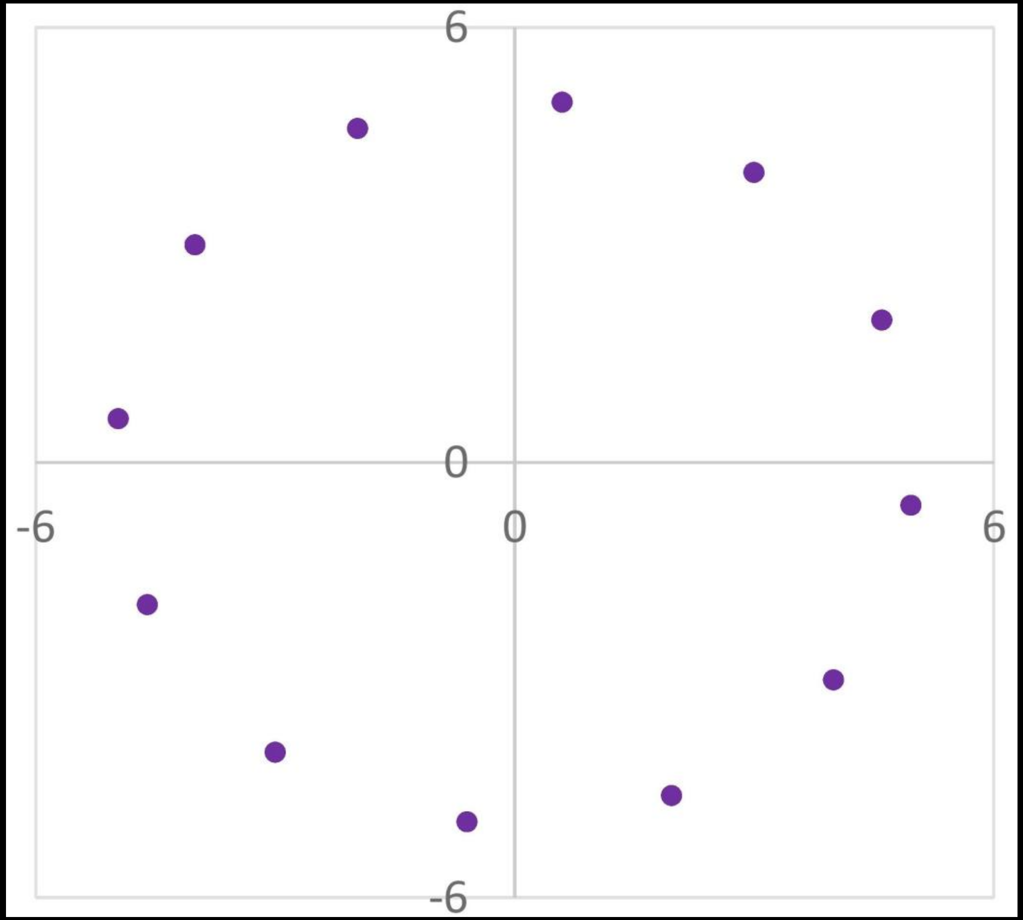
k	0	1	2	3	4	5	…	11	12	…
W_k R_k	3.00 4.00	4.60 1.96	4.96 -0.60	4.00 -3.00	1.96 -4.60	-0.60 -4.96	…	0.60 4.96	3.00 4.00	…

k	0	1	2	3	4	5	…	11	12	…
W_k R_k	3.00 4.00	4.60 1.96	4.96 −0.60	4.00 −3.00	1.96 −4.60	−0.60 −4.96	…	0.60 4.96	3.00 4.00	…



k	0	1	2	3	4	5	…	11	12	…
W_k R_k	3.00 4.00	4.60 1.96	4.96 −0.60	4.00 −3.00	1.96 −4.60	−0.60 −4.96	…	0.60 4.96	3.00 4.00	…





A number of issues and observations

- A matrix “does” something to a vector to get another vector.
- Linear transformations.
- Rotational matrices.
- Complex eigenvectors/values.

Variation of problem

$$x_{k+1} = Ax_k \quad \text{where} \quad A = \begin{bmatrix} 1.5 & 1 \\ 0.5 & 1 \end{bmatrix}, \; x_0 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}.$$

k	**0**	**1**	**2**	**3**	**4**	**5**	⋯	**50**	**51**	**52**
x_k	-3.0 7.0	2.5 5.5	9.3 6.8	20.6 11.4	42.3 21.7	85.2 42.8	…	$3.0E15$ $1.5E15$	$6.0E15$ $3.0E15$	$1.2E16$ $6.0E15$

- After several iterations, what is this vector that seems to be a multiple (2 ×) of the previous vector, after multiplying by A?
- Where did this multiple of 2 come from?
- Why in the previous example did vectors rotate around origin, but not here?
- (We are also seeing a dominant eigenvalue and corresponding eigenvector.)

Examples to motivate a chapter or section

- Chapters are typically built around a central theme. Why not give an example to motivate as much of the chapter as possible, without worrying too much about the details yet.
- We briefly revisit the example as we proceed through the chapter, as students are learning the ideas. We also revisit the example at the conclusion of the chapter, once students have learned all of the pertinent concepts and can fully work the problem.

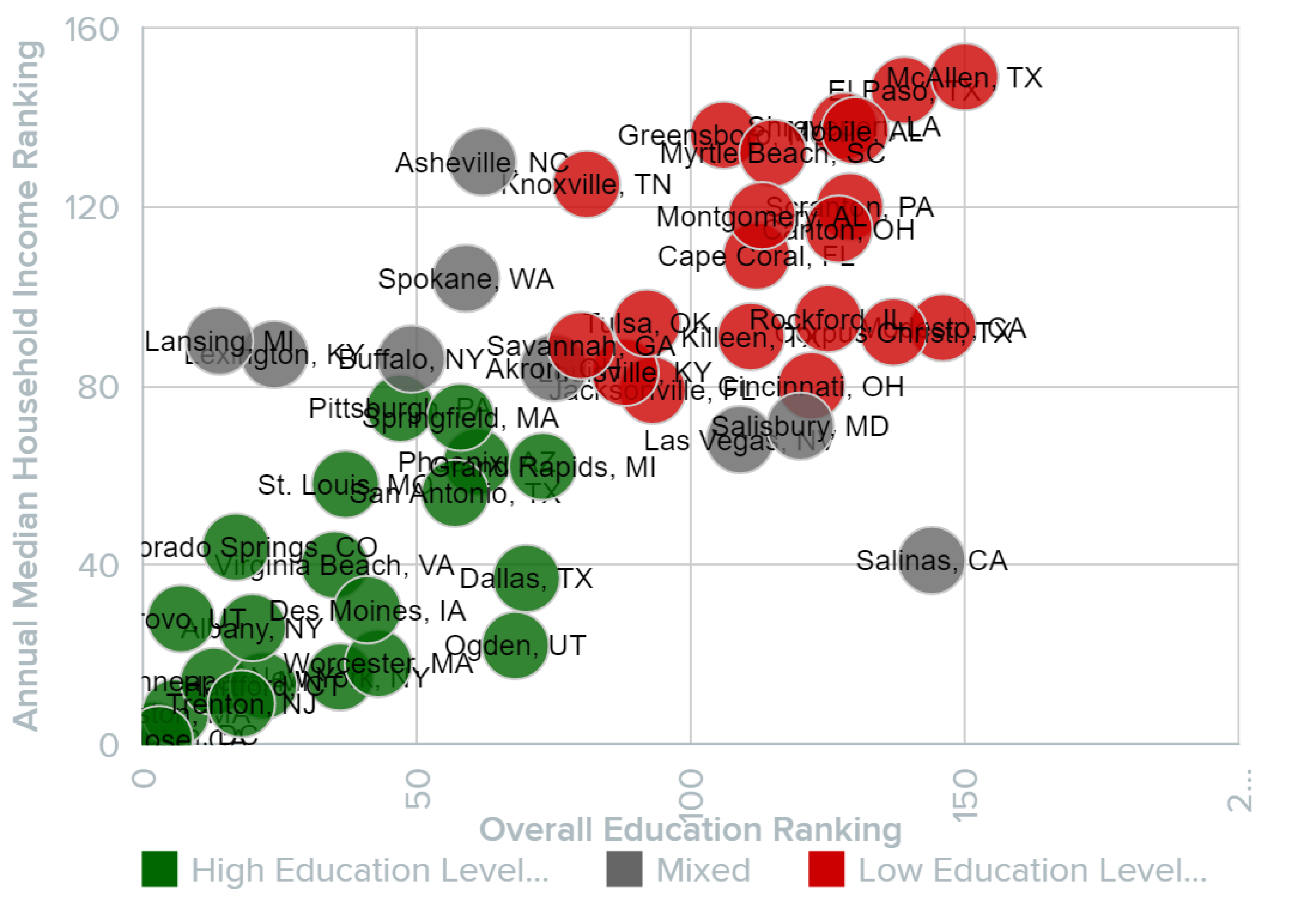
Least squares problem

- So we want to find the solution to a system of equations, for example,

$$\begin{aligned} n + \quad d &= \quad 6 \\ 5n - \quad d &= \quad 0 \\ 5n + 10d &= \quad 75 \end{aligned}, \quad \text{i.e.} \begin{bmatrix} 1 & 1 \\ 5 & -1 \\ 5 & 10 \end{bmatrix} \begin{bmatrix} n \\ d \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 75 \end{bmatrix}, \quad \text{i.e.} \; n \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} + d \begin{bmatrix} 1 \\ -1 \\ 10 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 75 \end{bmatrix}$$

(There are plenty of simple yet motivating real life examples of data fitting.)

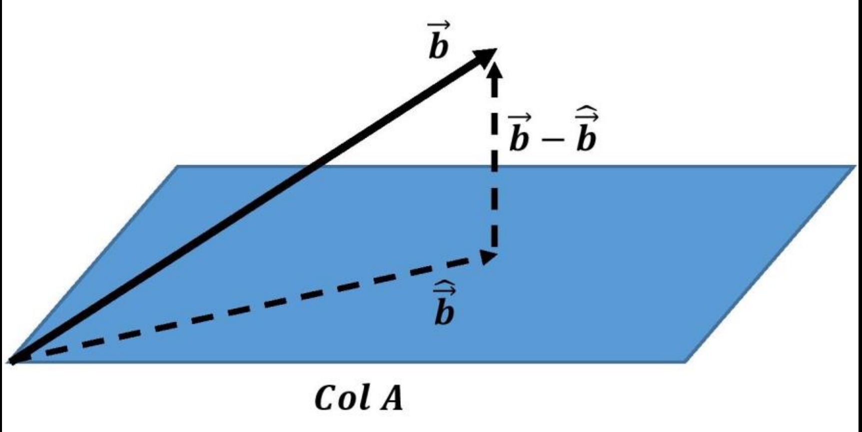
- In finding a solution to $A\vec{x} = \vec{b}$ we are trying to find how much $\vec{x}$ of each column in A is needed in order to build $\vec{b}$.
- What if $\vec{b}$ cannot be built from the columns of A?
  We build the vector $\hat{\vec{b}}$ that is close as possible to $\vec{b}$ using the columns of A.
- So how do we do this?



Annual Median Household Income Ranking
Overall Education Ranking
160
120
80
40
0
0
50
100
150
McAllen, TX
El Paso, TX
Mobile, AL
Greensboro
Myrtle Beach, SC
Asheville, NC
Knoxville, TN
Scranton, PA
Montgomery, AL
Canton, OH
Cape Coral, FL
Spokane, WA
Tulsa, OK
Rockford, IL
Killeen, TX
Lansing, MI
Lexington, KY
Buffalo, NY
Savannah, GA
Akron
Jacksonville, FL
Cincinnati, OH
Pittsburgh, PA
Springfield, MA
Salisbury, MD
Las Vegas, NV
Phoenix, AZ
Grand Rapids, MI
St. Louis, MO
San Antonio, TX
Colorado Springs, CO
Virginia Beach, VA
Salinas, CA
Dallas, TX
Provo, UT
Des Moines, IA
Albany, NY
Ogden, UT
Worcester, MA
Trenton, NJ
High Education Level...
Mixed
Low Education Level...

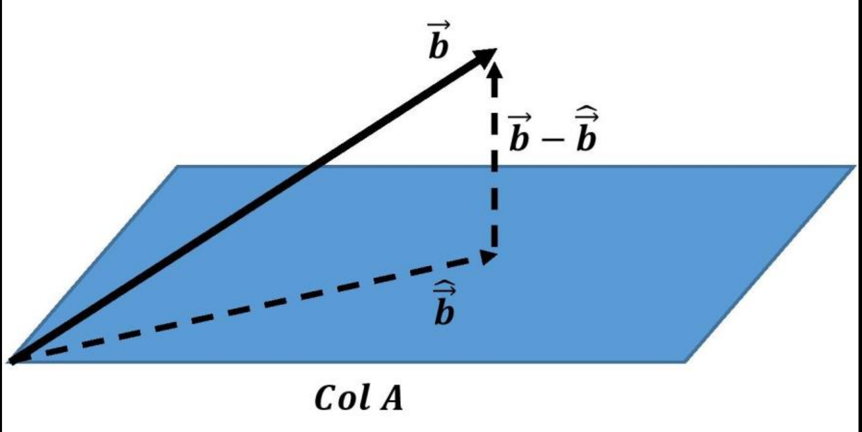
Least squares problem: visual explanation

- In finding a solution to $A\vec{x} = \vec{b}$ we are trying to find how much $\vec{x}$ of each column in A to build $\vec{b}$.
- What if $\vec{b}$ cannot be built from the columns of A? We build the vector $\hat{\vec{b}}$ that is close as possible to $\vec{b}$ using the columns of A.
- So how do we do this?



Least squares problem: visual explanation

- So how can we tell which vector $\hat{\vec{b}}$ is the one that is closest to $\vec{b}$? It's the vector "underneath" it.
- What do we mean by "underneath"?
  The difference $\vec{b} - \hat{\vec{b}}$ is orthogonal to $Col\ A$.
- So how can we tell when this is happening, especially in higher dimensions?
- And so on...



Least squares problem

- So we will have motivated several ideas that will come up in this chapter which culminates in solving least squares problems:
  - Properties of vectors, including orthogonality
  - Orthogonal projections
  - Projecting a vector onto the column space of a matrix
  - Least squares solutions
  - Extensions of these ideas to functions and other non-vectors
- The point is to give some reason, to help students see the need, as to why we want to learn about these various ideas.
- With some thought, I think we can do this with every chapter.

Meaning and Context

Meaning of homogeneous solution

- Solution $\vec{x}_h$ to $A\vec{x} = \vec{0}$.
- Where $A\vec{x}_p = \vec{b}$, is part of general solution $\vec{x}_p + c\vec{x}_h$ to $A\vec{x} = \vec{b}$.
- Both true, but not very motivating to most students.

Meaning of homogeneous solution: traffic flow

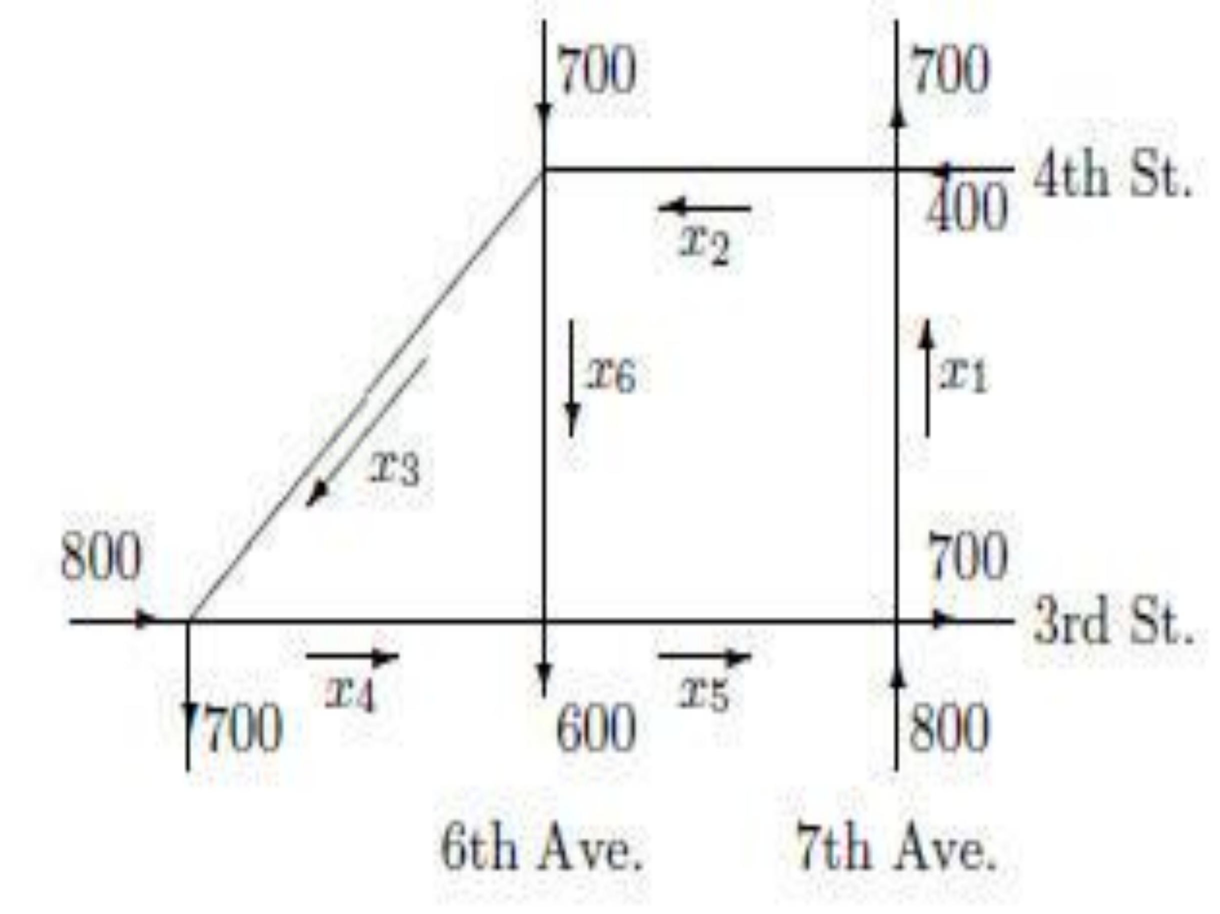
$x_1 + 400 = x_2 + 700$

$x_2 + 700 = x_3 + x_6$

$x_3 + 800 = x_4 + 700$

$x_4 + x_6 = x_5 + 600$

$x_5 + 800 = x_1 + 700$



Meaning of homogeneous solution: traffic flow

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 300 \\ -700 \\ -100 \\ 600 \\ 100 \end{bmatrix}$$

One possible description of solution:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 300 \\ 0 \\ 600 \\ 700 \\ 200 \\ 100 \end{bmatrix} + x_5 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Meaning of homogeneous solution: traffic flow

- The homogeneous solution is any linear combination of $\begin{bmatrix}1\\1\\1\\1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\0\\-1\\-1\\0\\1\end{bmatrix}$.

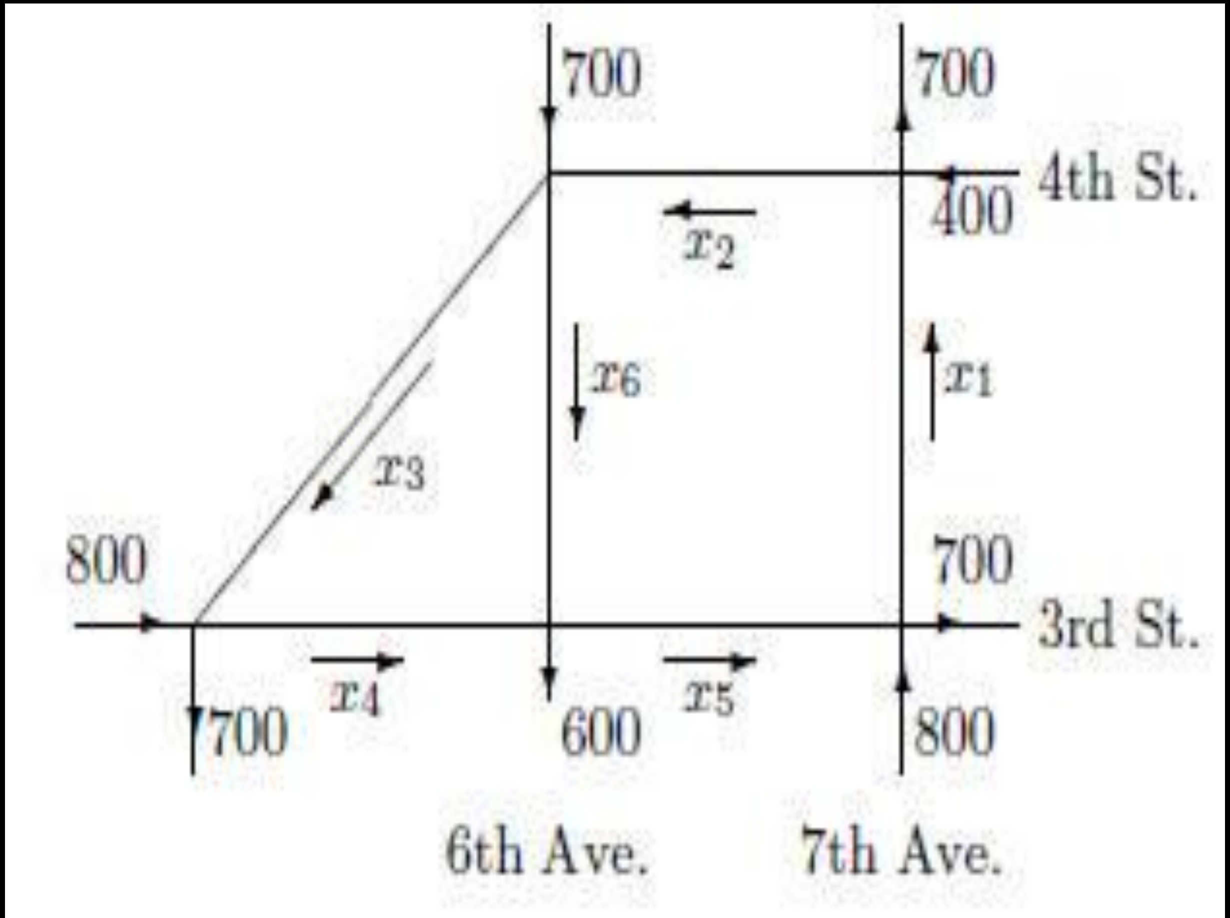
For example, their sum $\begin{bmatrix}1\\1\\0\\0\\1\\1\end{bmatrix}$ is also a homogenous solution.

Meaning of homogeneous solution: traffic flow

- So what's so special about

$$\begin{bmatrix}1\\1\\1\\1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\-1\\0\\1\end{bmatrix} \text{ and } \begin{bmatrix}1\\1\\0\\0\\1\\1\end{bmatrix}?$$

- (This is also a living example of a vector space.)



Meaning of homogeneous solution: investments

- Three investments x_1, x_2 and x_3 where

$$\begin{aligned} x_1 + \quad x_2 + \quad x_3 &= 10{,}000 \\ .05x_1 + .10x_2 + .25x_3 &= 1{,}000 \end{aligned}$$

which has general solution

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 10{,}000 \\ 0 \end{bmatrix} + c \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}.$$

- The homogeneous solution is the amounts each investment much change to still satisfy the given conditions.
- Simple, but again worth noting.

Summary

- I doubt any of today's examples are new to anyone.
- The point is, to make more effective and fuller use of examples and applications to motivate concepts that students will learn to give the students more motivation to learn those concepts.
- Also, more context and meaning for those concepts will enrich the learning. The concepts do not exist in a vacuum. They come from and are evident in many real situations and problems.

Thanks for your time

Thanks for your time