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

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

➤ Ax is the Matrix-by-vector product of A by 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)
- ✍ General rule or process?
- ✍ Cost?
- ✍ 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 $\begin{array}{rl} x_1 + 2x_2 & -x_3 = 4 \\ -5x_2 & +3x_3 = 1 \end{array}$ 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: $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

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

- $Ax = b$ is called a matrix equation.

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

- A is the coefficient matrix, b is the right-hand side

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

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

- 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.

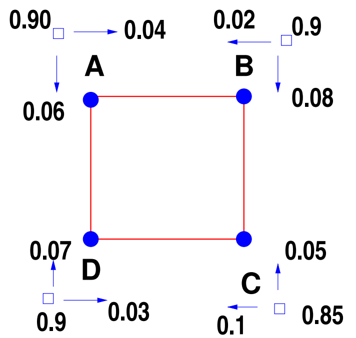
- So, it suffices to show (for an arbitrary matrix $\boldsymbol{A}$) that (1) is true iff (4) is true, i.e., that (1) and (4) are either both true or false.
- Given $\boldsymbol{b}$ in $\mathbb{R}^m$, we can row reduce the augmented matrix $[\boldsymbol{A}|\boldsymbol{b}]$ to reduced row echelon form $[\boldsymbol{U}|\boldsymbol{d}]$.
- 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 $[\boldsymbol{A}|\boldsymbol{b}]$ for a certain $\boldsymbol{b}$.
- The new system $\boldsymbol{Ax} = \boldsymbol{b}$ is also inconsistent, and (1) is false.

■

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.

$$\boldsymbol{x}^{(t)} = \begin{bmatrix} x_A^{(t)} \\ x_B^{(t)} \\ x_C^{(t)} \\ x_D^{(t)} \end{bmatrix}$$

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

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

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

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

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

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

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

Application: Google's Page rank

Note: Read this to prepare for HW2!

➤ Idea is to put order into the web by ranking pages by their importance..

➤ Install the google-toolbar on your laptop or computer

`http://toolbar.google.com/`

➤ Tells you how important a page is...

➤ Google uses this for searches..

➤ Updated regularly..

➤ Still a lot of mystery in what is in it..

Page-rank - explained

Main point: A page is important if it is pointed to by other important pages.

➤ Importance of your page (its PageRank) is determined by summing the page ranks of all pages which point to it.

➤ Weighting: If a page points to several other pages, then the weighting should be distributed proportionally.

➤ Imagine many tokens doing a random walk on this graph:

- (δ/n) chance to follow one of the n links on a page,
- $(1-\delta)$ chance to jump to a random page.
- What's the chance a token will land on each page?

➤ If `www.cs.umn.edu/~saad` points to 10 pages including yours, then you will get 1/10 of the credit of my page.

Page-Rank - definitions

➤ Build a 'Hyperlink' matrix H defined as follows

"every entry h_{ij} in column j is zero except when i is one of the links from j to i in which case $h_{ij} = 1/k_j$ where k_j = number of links from (j)"

➤ Defines a (possibly huge) Hyperlink matrix H

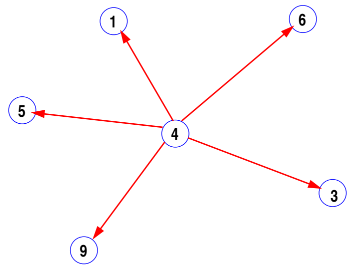
$$h_{ij} = \begin{cases} \frac{1}{k_j} & \text{if } j \text{ points to } i \\ 0 & \text{otherwise} \end{cases}$$

➤ Wll see to distinct cases: $\delta = 1$ (called undamped); $0 < \delta < 1$ (called damped)

➤ δ is a called a 'damping' parameter close to 1 – e.g. 0.85

Example: Here the 4th column of H consists of zeros except

$$h_{14} = 1/5;\ h_{34} = 1/5;$$
$$h_{64} = 1/5;\ h_{94} = 1/5$$
$$h_{54} = 1/5;$$



Simple case: $\delta = 1$

If token is at node j (with probability 1) at some stage, in the next stage it will jump to node i with probability h_{ij}.

➤ Case $\delta == 1$ will be very similar to the other Markov chain examples [population movement].

➤ Solved in exactly the same way.

- Issue: token can get stuck if a node has no outgoing links.

General case: $0 < \delta < 1$ Assumption: token has

- δ/k_j chance of jumping to one of the k_j links from j
- $1 - \delta$ chance to go to a random page

We wish to say next jump land in node i with a 'probability' of: $(1 - \delta) + \delta h_{ij}$ Dont add-up to 1

➤ Let $\rho_1, \rho_2, \cdots, \rho_n$ be n measures of importance for nodes $1, 2, \cdots, n$. [think of them as 'votes' or likelihoods of being visited]

➤ Google page-rank defines the ρ_i's by the following equation:

$$\rho_i = 1 - \delta + \delta \left[\frac{\rho_1}{k_1} + \frac{\rho_2}{k_2} + \cdots + \frac{\rho_n}{k_n} \right]$$

➤ ρ_i gets assigned a value that depends on the other ρ_j's

✍ Why is the above definition sensible?

➤ Let e be the vector of all ones (length n) and v the vector with components $\rho_1, \rho_2, \cdots, \rho_n$.

✍ Show that the above equation is equivalent to

$$v = (1 - \delta)e + \delta H v$$

✍ How would you solve the system?

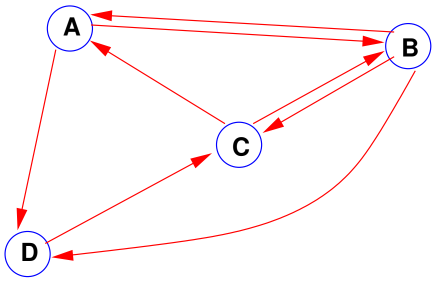
➤ Can show: Sum of all PageRanks == n: $\sum \rho_i = n$

✍ What is the 4×4 matrix H for the following case? [4 Nodes]

A points to B and D; B points to A, C, and D;
C points to A and B; D points to C;

Also: Determine the ρ_i's for this case when $\delta = 0.9$ (Matlab)

Solution:



	A	B	C	D
A	0	1/3	1/2	0
B	1/2	0	1/2	0
C	0	1/3	0	1
D	1/2	1/3	0	0

- Column- sums of H are $= 1$.
- If $\delta = .9$ then solving the linear system yields $v = \begin{bmatrix} 0.94144 \\ 1.05007 \\ 1.16982 \\ 0.83867 \end{bmatrix}$

The Google PageRank algorithm

- As one can imagine H can be huge so solving the linear system by GE is not practical.

Alternative: following iterative algorithm

Algorithm (PageRank)

1. Select initial vector v ($v \geq 0$)
2. For i=1:maxitr
3 $\quad v := (1 - \delta)e + \delta H v$
4. end

Do a few steps of this algorithm for previous example