Linear Algebra

Linear Algebra notes
MT 1004

Muhammad Shahid Ashraf

Notes

Gram-Schmidt and QR process

Notes

Orthogonal sets

The vectors $\mathbf{u} = (1, 1)$ and $\mathbf{v} = (1, -1)$ are orthogonal

$$U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

Notes

Let P_2 have the inner product $\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} p(x)q(x)\,dx$

and let $\mathbf{p} = x$ and $\mathbf{q} = x^2$. Then

$$\|\mathbf{p}\| = \langle \mathbf{p}, \mathbf{p} \rangle^{1/2} = \left[\int_{-1}^{1} xx\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^2\,dx\right]^{1/2} = \sqrt{\frac{2}{3}}$$

$$\|\mathbf{q}\| = \langle \mathbf{q}, \mathbf{q} \rangle^{1/2} = \left[\int_{-1}^{1} x^2x^2\,dx\right]^{1/2} = \left[\int_{-1}^{1} x^4\,dx\right]^{1/2} = \sqrt{\frac{2}{5}}$$

$$\langle \mathbf{p}, \mathbf{q} \rangle = \int_{-1}^{1} xx^2\,dx = \int_{-1}^{1} x^3\,dx = 0$$

Because $\langle \mathbf{p}, \mathbf{q} \rangle = 0$, the vectors $\mathbf{p} = x$ and $\mathbf{q} = x^2$ are orthogonal

$$\|\mathbf{p} + \mathbf{q}\|^2 = \left(\sqrt{\frac{2}{3}}\right)^2 + \left(\sqrt{\frac{2}{5}}\right)^2 = \frac{2}{3} + \frac{2}{5} = \frac{16}{15}$$

$$\|\mathbf{p} + \mathbf{q}\|^2 = \langle \mathbf{p} + \mathbf{q}, \mathbf{p} + \mathbf{q} \rangle = \int_{-1}^{1} (x + x^2)(x + x^2)\,dx$$

$$= \int_{-1}^{1} x^2\,dx + 2\int_{-1}^{1} x^3\,dx + \int_{-1}^{1} x^4\,dx = \frac{2}{3} + 0 + \frac{2}{5} = \frac{16}{15}$$

Notes

Orthogonal set

Let $\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = (1,0,1)$, $\mathbf{v}_3 = (1,0,-1)$
and assume that R^3 has the Euclidean inner product.
$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set since $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \langle \mathbf{v}_1, \mathbf{v}_3\rangle = \langle \mathbf{v}_2, \mathbf{v}_3\rangle = 0$.

$$\|\mathbf{v}_1\| = 1, \quad \|\mathbf{v}_2\| = \sqrt{2}, \quad \|\mathbf{v}_3\| = \sqrt{2}$$

Consequently, normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (0,1,0), \quad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

$$\langle \mathbf{u}_1, \mathbf{u}_2\rangle = \langle \mathbf{u}_1, \mathbf{u}_3\rangle = \langle \mathbf{u}_2, \mathbf{u}_3\rangle = 0 \quad \text{and} \quad \|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \|\mathbf{u}_3\| = 1$$

Notes

Orthogonal projection

Theorem Let V be an inner product space and V_0 be a finite-dimensional subspace of V. Then any vector $\mathbf{x} \in V$ is uniquely represented as $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in V_0$ and $\mathbf{o} \perp V_0$.

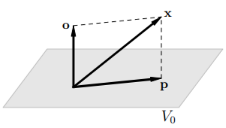
The component $\mathbf{p}$ is the **orthogonal projection** of the vector $\mathbf{x}$ onto the subspace V_0. We have

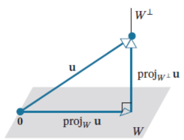
$$\|\mathbf{o}\| = \|\mathbf{x} - \mathbf{p}\| = \min_{\mathbf{v} \in V_0} \|\mathbf{x} - \mathbf{v}\|.$$

That is, the distance from $\mathbf{x}$ to the subspace V_0 is $\|\mathbf{o}\|$.

Notes

Orthogonal projection





Notes

Orthogonal projection

- Let $\mathbb{R}^3$ have the Euclidean inner product, and let W be the subspace spanned by the orthonormal vectors

$$v_1 = (0,1,0) and v_2 = (-45, 0, 35).$$

- Find the orthogonal projection of $u = (1,1,1)$ on W

$$proj_W u = < u, v_1 > v_1 + < u, v_2 > v_2$$
$$= (1)(0,1,0) + (\frac{-1}{5})(\frac{-4}{5}, 0, \frac{3}{5})$$
$$= (\frac{4}{25}, 1, \frac{3}{25})$$

Notes

Let V be a vector space with an inner product. Suppose $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is a basis for V. Let

$$\mathbf{v}_1 = \mathbf{x}_1,$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1,$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2,$$

....

$$\mathbf{v}_n = \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle}\mathbf{v}_{n-1}.$$

Then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V.

Notes

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$\vdots$

(continue for r steps)

Notes

Assume that the vector space R^3 has the Euclidean inner product. Apply the Gram–Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1, 1, 1), \quad \mathbf{u}_2 = (0, 1, 1), \quad \mathbf{u}_3 = (0, 0, 1)$$

into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

Notes

Step 1. $\mathbf{v}_1 = \mathbf{u}_1 = (1, 1, 1)$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{W_1}\mathbf{u}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

$$= (0, 1, 1) - \frac{2}{3}(1, 1, 1) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

$$= (0, 0, 1) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$= \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad \mathbf{v}_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

form an orthogonal basis for R^3.

The norms of these vectors are

$$\|\mathbf{v}_1\| = \sqrt{3}, \quad \|\mathbf{v}_2\| = \frac{\sqrt{6}}{3}, \quad \|\mathbf{v}_3\| = \frac{1}{\sqrt{2}}$$

so an orthonormal basis for R^3 is

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right),$$

$$\mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right),$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Notes

QR-Decomposition

- If A is an m × n matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

Notes

QR-Decomposition

Find a QR-decomposition of

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

The column vectors of A are

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$R = \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{q}_1 \rangle & \langle \mathbf{u}_2, \mathbf{q}_1 \rangle & \langle \mathbf{u}_3, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{u}_2, \mathbf{q}_2 \rangle & \langle \mathbf{u}_3, \mathbf{q}_2 \rangle \\ 0 & 0 & \langle \mathbf{u}_3, \mathbf{q}_3 \rangle \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Notes

Page Rank and Markov Chains

Notes

Markov Chains

- Consider a system that is, at any one time, in one and only one of a finite number of states.
- For example,
  - the weather in a certain area is either rainy or dry;
  - a person is either a smoker or a nonsmoker;
  - a person either goes or does not go to college;
  - we live in an urban,suburban, or rural area;
  - we are in the lower, middle, or upper income brackets;
  - we buy a Chevrolet, Ford, or other make of car.
- As time goes by, the system may move from one state to another,
- we assume that the state of the system is observed at fixed time intervals
- we know the present state of the system and we wish to know the state at the next, or some other future observation period.

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

- A Markov process is a process in which
  (1) The probability of the system being in a particular state at a given point in time depends only on its state at the immediately preceding observation period.
  (2) The probabilities are constant over time
  (3) The set of possible states/outcomes is finite.
- Suppose a system has n possible states. For each $i = 1, 2, 3 \cdots, n, j = 1, 2, 3 \cdots, n$, let p_{ij} be the probability that if part of the system is in state j at the current time period, then it will be in state i at the next.
- A transition probability is an entry p_{ij} in a stochastic/transition matrix. That is, it is a number representing the chance that something in state j right now will be in state i at the next time interval.

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

- Coca-cola is testing a new diet version of their best-selling soft drink, in a small town in California. They poll shoppers once per month to determine what customers think of the new product. Suppose they find that every month, $\frac{1}{3}$ of the people who bought the diet version decide to switch back to regular, and $\frac{1}{2}$ the people who bought diet decide to switch to the new diet version. Let D denote diet soda buyers, and let R be regular soda buyers. Then the transition matrix of this Markov process is

$$P = \begin{bmatrix} \frac{2}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix}$$

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

A market research organization is studying a large group of caffeine addicts who buy a can of coffee each week. It is found that 50% of those presently using Starbuck's will again buy Starbuck's brand next week, 25% will switch to Peet's, and 25% will switch to some brand. Of those buying Peet's now, 30% will again buy Peet's next week, 60% will switch to Starbuck's, and 10% will switch to another brand. Of those using another brand now, 40% will switch to Starbuck's and 30% will switch to Peet's in the next week. Let S, P, and O denote Starbuck's, Peet's and Other, respectively. The probability that a person presently using S will switch to P is 0.25, the probability that a person presently using P will again buy P is 0.3, and so on. Thus, the transition matrix of this Markov process is

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

A market research organization is studying a large group of caffeine addicts who buy a can of coffee each week. It is found that 50% of those presently using Starbuck's will again buy Starbuck's brand next week, 25% will switch to Peet's, and 25% will switch to some brand. Of those buying Peet's now, 30% will again buy Peet's next week, 60% will switch to Starbuck's, and 10% will switch to another brand. Of those using another brand now, 40% will switch to Starbuck's and 30% will switch to Peet's in the next week. Let S, P, and O denote Starbuck's, Peet's and Other, respectively. The probability that a person presently using S will switch to P is 0.25, the probability that a person presently using P will again buy P is 0.3, and so on. Thus, the transition matrix of this Markov process is

$$\begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix}$$

Notes

Markov Chains

A **probability vector** is a vector

$$\bar{\mathbf{x}} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}$$

(1) whose entries p_i are between 0 and 1: $0 \leqslant p_i \leqslant 1$, and
(2) whose entries p_i sum to 1: $p_1 + p_2 + \ldots + p_n = \sum_{i=1}^{n} p_i = 1$

Each column of the coffee transition matrix is a probability vector:

$$\bar{\mathbf{x}}_S = \begin{bmatrix} 0.50 \\ 0.25 \\ 0.25 \end{bmatrix} \qquad \bar{\mathbf{x}}_P = \begin{bmatrix} 0.60 \\ 0.30 \\ 0.10 \end{bmatrix} \qquad \bar{\mathbf{x}}_O = \begin{bmatrix} 0.40 \\ 0.30 \\ 0.30 \end{bmatrix}$$

Notes

Markov Chains

The **state vector** of a Markov process at step k is a probability vector

$$\bar{\mathbf{x}}^{(k)} = \begin{bmatrix} p_1^{(k)} \\ p_2^{(k)} \\ \vdots \\ p_n^{(k)} \end{bmatrix}$$

which gives the breakdown of the population at step k.
The state vector $\bar{\mathbf{x}}^{(0)}$ is the **initial state vector**.

Notes

Markov Chains

Let's consider the Coca-cola example Suppose that when we begin market observations, the initial state vector is

$$\bar{\mathbf{x}}^{(0)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

because Coca-Cola is giving away free samples of their product to everyone. (This vector corresponds to 100% of the people getting the diet version.) Then on month 1 (one month after the product launch), the state vector is

$$\bar{\mathbf{x}}^{(1)} = P\bar{\mathbf{x}}^{(0)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

That is, one-third of the people switched back to the regular version immediately.

Notes

Markov Chains

$$\bar{\mathbf{x}}^{(2)} = P\bar{\mathbf{x}}^{(1)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3}\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{1}{3} \\ \frac{1}{3}\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{11}{18} \\ \frac{7}{18} \end{bmatrix} \approx \begin{bmatrix} 0.611 \\ 0.389 \end{bmatrix}$$

$$\bar{\mathbf{x}}^{(3)} = P\bar{\mathbf{x}}^{(2)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{11}{18} \\ \frac{7}{18} \end{bmatrix} = \begin{bmatrix} \frac{2}{3}\cdot\frac{11}{18} + \frac{1}{2}\cdot\frac{7}{18} \\ \frac{1}{3}\cdot\frac{11}{18} + \frac{1}{2}\cdot\frac{7}{18} \end{bmatrix} = \begin{bmatrix} \frac{65}{108} \\ \frac{43}{108} \end{bmatrix} \approx \begin{bmatrix} 0.602 \\ 0.398 \end{bmatrix}$$

$$\bar{\mathbf{x}}^{(4)} = P\bar{\mathbf{x}}^{(3)} = \begin{bmatrix} \frac{2}{3} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{65}{108} \\ \frac{43}{108} \end{bmatrix} = \begin{bmatrix} \frac{2}{3}\cdot\frac{65}{108} + \frac{1}{2}\cdot\frac{43}{108} \\ \frac{1}{3}\cdot\frac{65}{108} + \frac{1}{2}\cdot\frac{43}{108} \end{bmatrix} = \begin{bmatrix} \frac{389}{648} \\ \frac{259}{648} \end{bmatrix} \approx \begin{bmatrix} 0.600 \\ 0.400 \end{bmatrix}$$

From the fourth day on, the state vector of the system only gets closer to [0.60 0.40]. A very practical application of this technique might be to answer the question, "If we want our new product to eventually retain x% of the market share, what portion of the population must we initially introduce to the product?" In other words, "How much do we need to give away now in order to make a profit later?"

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

Consider the coffee example again. Suppose that when the survey begins, we find that Starbuck's has 20% of the market, Peet's has 20% of the market, and the other brands have 60% of the market. Then the initial state vector is

$$\bar{x}^{(0)} = \begin{bmatrix} 0.2 \\ 0.2 \\ 0.6 \end{bmatrix} .$$

The state vector after the first week is

$$\bar{x}^{(1)} = P\bar{x}^{(0)} = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} \begin{bmatrix} 0.20 \\ 0.20 \\ 0.60 \end{bmatrix} = \begin{bmatrix} 0.4600 \\ 0.2900 \\ 0.2500 \end{bmatrix}$$

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

Similarly,

$$\bar{x}^{(2)} = P\bar{x}^{(1)} = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} \begin{bmatrix} 0.4600 \\ 0.2900 \\ 0.2500 \end{bmatrix} = \begin{bmatrix} 0.5040 \\ 0.2770 \\ 0.2190 \end{bmatrix}$$

$$\bar{x}^{(3)} = P\bar{x}^{(2)} = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} \begin{bmatrix} 0.5040 \\ 0.2770 \\ 0.2190 \end{bmatrix} = \begin{bmatrix} 0.5058 \\ 0.2748 \\ 0.2194 \end{bmatrix}$$

$$\bar{x}^{(4)} = P\bar{x}^{(3)} = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} \begin{bmatrix} 0.5058 \\ 0.2748 \\ 0.2194 \end{bmatrix} = \begin{bmatrix} 0.5055 \\ 0.2747 \\ 0.2198 \end{bmatrix}$$

$$\bar{x}^{(5)} = P\bar{x}^{(4)} = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} \begin{bmatrix} 0.5055 \\ 0.2747 \\ 0.2198 \end{bmatrix} = \begin{bmatrix} 0.5055 \\ 0.2747 \\ 0.2198 \end{bmatrix}$$

So as k increases (that is, as time passes), we see that the state vector approaches the fixed vector

$$\bar{x} = \begin{bmatrix} 0.5055 \\ 0.2747 \\ 0.2198 \end{bmatrix} .$$

This means that in the long run, Starbuck's will command about 51% of the market, Peet's will retain about 27%, and the other brands will have about 22%.

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

The following example shows that not every Markov process reaches an equilibrium. Let

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \bar{x}^{(0)} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{2} \end{bmatrix}$$

Then

$$\bar{x}^{(1)} = P\bar{x}^{(0)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 + \frac{2}{3} \\ \frac{1}{3} + 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$\bar{x}^{(2)} = P\bar{x}^{(1)} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \text{and} \quad \bar{x}^{(3)} = P\bar{x}^{(2)} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Thus the state vector oscillates between the vectors

$$\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

and does not converge to a fixed vector.

Notes

Markov Chains

NATIONAL UNIVERSITY of Computer & Emerging Sciences

The second method is:
(a) Solve the homogeneous system $(I_n - P)\bar{b} = \mathbf{0}$.
(b) From the infinitely many solutions obtained this way, determine the unique solution whose components satisfy $b_1 + b_2 + \ldots + b_n = 1$.

Now let's return to the coffee example. For

$$P = \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix},$$

we have

$$(I_n - P) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.50 & 0.60 & 0.40 \\ 0.25 & 0.30 & 0.30 \\ 0.25 & 0.10 & 0.30 \end{bmatrix} = \begin{bmatrix} 0.50 & -0.60 & -0.40 \\ -0.25 & 0.70 & -0.30 \\ -0.25 & -0.10 & 0.70 \end{bmatrix}$$

This gives the homogeneous system

$$\begin{bmatrix} 0.50 & -0.60 & -0.40 \\ -0.25 & 0.70 & -0.30 \\ -0.25 & -0.10 & 0.70 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Notes

The Singular Value Decomposition

Notes

The Singular Value Decomposition

NATIONAL UNIVERSITY of Computer & Emerging Sciences

- Let A be an $m \times n$ matrix with rank r.
- Then there exists an $m \times n$ matrix Σ for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq, \ldots \geq \sigma_r$ 0,
- and there exist an m x m orthogonal matrix U and an n x n orthogonal matrix V such that

$$A = U\Sigma V^T$$

Notes

The Singular Value Decomposition

NATIONAL UNIVERSITY of Computer & Emerging Sciences

- Singular Value Decomposition is a decomposition $A = U * \Sigma * V^T$, where U and V are unitary matrices $(U * U^T = U^T * U = I$ and $V * V^T = V^T * V = I)$,
- Σ a diagonal matrix with non-negative entries.

$$\begin{aligned} A^T * A * V &= (U * \Sigma * V^T)^T * (U * \Sigma * V^T) * V \\ &= V * \Sigma^T * U^T * U * \Sigma * V^T * V \\ &= V * \Sigma^T * (U^T * U) * \Sigma * (V^T * V) \\ &= V * (\Sigma^T * \Sigma) \end{aligned}$$

- so $A^T * A * v_i = v_i * \sigma_i^2$ (where v_i is a column vector of V), which means v_i is an eigenvector of $A^T * A$ corresponding to an eigenvalue σ_i^2

Notes

The Singular Value Decomposition

NATIONAL UNIVERSITY of Computer & Emerging Sciences

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \text{ and } AA^T = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix}$$

$$\det(A^T A - \lambda I) = (5-\lambda)^2 - 25 = 0 \Rightarrow \lambda = 0 \text{ or } \lambda = 10$$

$$\text{For } \lambda_1 = 0, \mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \text{ For } \lambda_2 = 10, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \Rightarrow V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\det(AA^T - \lambda I) = (8-\lambda)(2-\lambda) - 16 = 0 \Rightarrow \lambda = 0 \text{ or } \lambda = 10$$

$$\text{For } \lambda_1 = 0, \mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}. \text{ For } \lambda_2 = 10, \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \Rightarrow U = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

$$\text{Let } \Sigma = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{10} \end{bmatrix} \Rightarrow U\Sigma V^{-1} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & \sqrt{10} \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right)^{-1}$$

$$= \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = A$$

Notes

The Singular Value Decomposition

- Find the singular value decomposition of

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$$

Notes

Find an orthogonal diagonalization of A^TA

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

The eigenvalues of A^TA are $\lambda_1 = 360, \lambda_2 = 90$, and $\lambda_3 = 0$. Corresponding unit eigei vectors are, respectively,

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

Set up V and Σ.

the singular values of A are

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = 0$$

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix} \qquad V = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

$$\Sigma = [\, D \quad 0 \,] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Construct U.

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{6\sqrt{10}} \begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} \quad \mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{3\sqrt{10}} \begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

$$A = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

Notes

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}. \quad A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

The eigenvalues of A^TA are 18 and 0

with corresponding unit eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

These unit vectors form the columns of $V = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \,] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

The singular values are $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = 0$.

The matrix Σ is the same size as A, with D in its upper left corner:

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

To construct U, first construct $A\mathbf{v}_1$ and $A\mathbf{v}_2$: $A\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix}, \quad A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\|A\mathbf{v}_1\| = \sigma_1 = 3\sqrt{2}. \quad \|A\mathbf{v}_2\| = \sigma_2 = 0.$

$\mathbf{u}_1 = \frac{1}{3\sqrt{2}} A\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ $\mathbf{u}_1^T\mathbf{x} = 0$ A basis for the solution set of this equation is

$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

Check that $\mathbf{w}_1$ and $\mathbf{w}_2$ are each orthogonal to $\mathbf{u}_1$

Apply the Gram–Schmidt process with normalizations) to $\{\mathbf{w}_1, \mathbf{w}_2\}$, and obtain

$$\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} \quad \mathbf{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix},$$

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Notes

Notes