Matris Çarpımı

T.C. Trakya Üniversitesi
Mühendislik Fakültesi
Elektrik - Elektronik Mühendisliği Bölümü
Kontrol Anabilim Dalı

Dr. Öğr. Üyesi Işık İlber Sırmatel
sirmatel.github.io

Kaynak (source)

Lecture Slides for Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares.
Stephen Boyd, Lieven Vandenberghe

Konu listesi

1. Tanım ve notasyon
2. Doğrusal fonksiyonların bileşkesi
3. Matris üsleri
4. QR ayrıştırması

Bölüm 1

Tanım ve notasyon

Matris çarpımı

- $m \times p$ matris A ile $p \times n$ matris B çarpılarak $m \times n$ matris C elde edilir:

  $$C_{ij} = \sum_{k=1} p A_{ik} B_{kj} = A_{i1} B_{1j} + \cdots + A_{ip} B_{pj}$$

  ($i = 1, \ldots, m,\ j = 1, \ldots, n$ için)
- örnek:

  $$\begin{bmatrix} -1.5 & 3 & 2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3.5 & -4.5 \\ -1 & 1 \end{bmatrix}$$

Matris çarpımının özel halleri

- skaler-vektör çarpımı (skaler sağda): $x\alpha$
- iç çarpım: a^Tb
- matris-vektör çarpımı: Ax
- m-vektör a ile n-vektör b'nin dış çarpımı (sonuç $m \times n$ matris):

$$ab^T = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_mb_1 & a_mb_2 & \cdots & a_mb_n \end{bmatrix}$$

Matris çarpımının özellikleri

- $(AB)C = A(BC)$ (dolayısıyla her ikisi de ABC olarak yazılabilir)
- $A(B+C) = AB + AC$
- $(AB)^T = B^T A^T$
- $AI = A$ ve $IA = A$
- $AB = BA$ genel olarak geçerli değildir

Blok matrislerin çarpımı

blok matrisler aynı yöntemle çarpılabilir, örneğin:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ F & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CD + DG & CF + DH \end{bmatrix}$$

(bütün çarpımların olanaklı olduğunu varsayarak)

Sütun yorumu

- B matrisinin sütunlarına b_i diyelim:

$$B = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$$

- bu durumda A ve B'nin çarpımı

$$\begin{aligned} AB &= A \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \\ &= \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_n \end{bmatrix} \end{aligned}$$

  şeklinde olur
- dolayısıyla AB, A ile B'nin sütunlarının yığın çarpımıdır

Çoklu doğrusal denklem takımları

- katsayı matrisleri aynı olan ($m \times n$ matris A) k adet doğrusal denklem takımını ele alalım:

$$Ax_i = b_i, \quad i = 1, \ldots, k$$

- kompakt matris formunda yazarsak: $AX = B$
- burada $X = \begin{bmatrix} x_1 & \cdots & x_k \end{bmatrix}$, $B = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix}$

İç çarpım yorumu

- A'nın satırlarına a_i^T, B'nin sütunlarına b_j dersek:

$$AB = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_n \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_n \end{bmatrix}$$

- dolayısıyla, matris çarpımı A'nın satırları ile B'nin sütunlarının bütün iç çarpımlarının bir matris formunda düzenlenmesidir

Gram matrisi

- A, sütunları $a_1, \ldots, a_n$ olan $m \times n$ matris olsun
- A'nın Gram matrisi:

$$G = A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{bmatrix}$$

- Gram matrisi A'nın sütunlarının bütün iç çarpımlarını verir
- örnek: $G = A^T A = I$ ise A'nın sütunları birim dikgendir

Karmaşıklık

- $C_{ij} = (AB)_{ij}$'nin hesabı için iki p-vektörün iç çarpımı gerekir
- matris çarpımının toplam maliyeti: $(mn)(2p) = 2mnp$ flop
- iki 1000×1000 matrisin çarpımı için 2 milyar flop gerekir; bu işlem modern bilgisayarlada 1 saniyeden çok daha kısa sürede yapılabilir

Bölüm 2

Doğrusal fonksiyonların bileşkesi

Doğrusal fonksiyonların bileşkesi

- A $m \times p$ matris, B $p \times n$ matris olsun
- $f : \mathbb{R}^p \to \mathbb{R}^m$ ve $g : \mathbb{R}^n \to \mathbb{R}^p$ fonksiyonlarını

  $$f(u) = Au, \qquad g(v) = Bv$$

  şeklinde tanımlayalım
- f ve g doğrusal fonksiyonlardır
- f ve g'nin bileşkesi (*composition*): $h : \mathbb{R}^n \to \mathbb{R}^m$, $h(x) = f(g(x))$
- $h(x)$'yi şu şekilde yazabiliriz:

  $$h(x) = f(g(x)) = A(Bx) = (AB)x$$

- doğrusal fonksiyonların bileşkesi doğrusaldır
- ilgili matris, fonksiyonların matrislerinin çarpımıdır

İkinci fark matrisi

- D_n $(n-1) \times n$ fark matrisi:

$$D_n x = \begin{bmatrix} x_2 - x_1 & \dots & x_n - x_{n-1} \end{bmatrix}^T$$

- D_{n-1} $(n-2) \times (n-1)$ fark matrisi:

$$D_{n-1} y = \begin{bmatrix} y_2 - y_1 & \dots & y_{n-1} - y_{n-2} \end{bmatrix}^T$$

- $\Delta = D_{n-1} D_n$ $(n-2) \times n$ ikinci fark matrisi:

$$\Delta x = \begin{bmatrix} x_1 - 2x_2 + x_3 & x_2 - 2x_3 + x_4 & \dots & x_{n-2} - 2x_{n-1} + x_n \end{bmatrix}^T$$

- $n = 5$ için $\Delta = D_{n-1} D_n$:

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

Bölüm 3

Matris üsleri

Matris üsleri

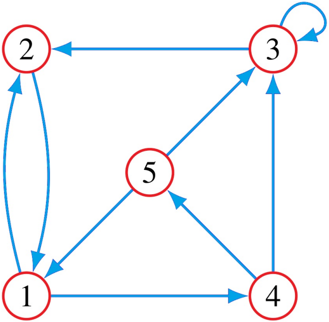
- kare matris A için A^2 AA anlamına gelir; daha büyük üsler için de aynıdır
- $A^0 = I$ konvansiyonuyla $A^k A^l = A^{k+l}$ yazabiliriz
- negatif üsleri daha sonra inceleyeceğiz; kesirli üsler bu dersin kapsamı dışında

Yönlü çizge

- bir yönlü çizgenin bitişiklik (*adjacency*) matrisi A

$$A_{ij} = \begin{cases} 1 & \text{düğüm } j\text{'den düğüm } i\text{'ye giden bir ayrıt var} \\ 0 & \text{aksi halde} \end{cases}$$

  şeklinde tanımlı bir $n \times n$ matristir
- örnek:



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

Yönlü çizgede yolaklar

- bitişiklik matrisinin karesi:

$$(A^2)_{ij} = \sum_{k=1}^{n} A_{ik} A_{kj}$$

- $(A^2)_{ij}$, düğüm j'den düğüm i'ye giden ve uzunluğu iki olan yolakların (*path*) sayısıdır
- önceki sayfadaki örnek için:

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

  örneğin, düğüm 4'ten düğüm 3'e (ayrıt 3 ve ayrıt 5 üzerinden) giden iki yolak vardır
- daha genel olarak, $(A^l)_{ij}$ düğüm j'den düğüm i'ye giden ve uzunluğu l olan yolakların sayısıdır

Bölüm 4

QR ayrıştırması

Matris notasyonuyla Gram-Schmidt

- $n \times k$ matris A'nın sütunları $a_1, \ldots, a_k$ için Gram-Schmidt algoritmasını çalıştırdığımızı varsayalım
- sütunlar doğrusal bağımsız ise, birim dikgen vektörler $q_1, \ldots, q_k$ elde edilir
- sütunları $q_1, \ldots, q_k$ olan $n \times k$ matris Q tanımlayalım
- $Q^T Q = I$ (yani, Q'nun sütunları birim dikgen)
- Gram-Schmidt algoritmasından:

  $$\begin{aligned} a_i &= (q_1^T a_i) q_1 + \cdots + (q_{i-1}^T a_i) q_{i-1} + \|\tilde{q}_i\| q_i \\ &= R_{1i} q_1 + \cdots + R_{ii} q_i \end{aligned}$$

  (burada $i < j$ için $R_{ij} = q_i^T a_j$ ve $R_{ii} = \|\tilde{q}_i\|$)
- $i > j$ için $R_{ij} = 0$ tanımıyla $A = QR$ elde edilir
- R üst üçgen ve köşegen elemanları pozitif

QR ayrıştırması

- $A = QR$'ye A'nın QR ayrıştırması denir
- çarpanlar (yani, Q ve R) şu şartları sağlar:
  - $Q^T Q = I$
  - R üst üçgen ve köşegen elemanları pozitif
- QR ayrıştırması Gram-Schmidt algoritması (veya bazı çeşitleri) kullanılarak hesaplanabilir
- çok sayıda kullanımı vardır. bunları ilerleyen kısımlarda inceleyeceğiz