Matrisler (temel kavramlar)

Dr. Öğr. Üyesi Işık İlber Sırmatel

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

Kaynak (source)

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

Konu listesi

Bölüm 1

Tanım ve notasyon

Matrisler

- **matris** dikdörtgen şeklinde bir sayı dizilimidir (*array*) ve (örneğin)
$$\begin{bmatrix} 0 & 1 & -2.3 & 0.1 \\ 1.3 & 4 & -0.1 & 0 \\ 4.1 & -1 & 0 & 1.7 \end{bmatrix}$$
şeklinde yazılır
- boyutu (satır (*row*) boyutu)$\times$(sütun (*column*) boyutu) şeklinde verilir (örneğin, yukarıdaki matris 3×4)
- dizilimdeki sayılara matrisin **eleman**ları denir
- elemanlara öğe (*entry*) veya katsayı (*coefficient*) da denir
- B isimli bir matrisin satır i ve sütun j'deki elemanı (yani, i, j elemanı) B_{ij} ile gösterilir
- i satır indisi, j ise sütun indisidir; indisler 1'den başlar
- aynı boyutlu iki matris A ve B'nin bütün karşılıklı elemanları (yani, A_{ij} ve B_{ij}) eşitse matrisler **eşit**tir; bu durum $A = B$ ile gösterilir

Matris şekilleri

bir $m \times n$ A matrisine

- $m > n$ ise uzun (*tall*) matris
- $m < n$ ise geniş (*wide*) matris
- $m = n$ ise kare (*square*) matris

denir

Satır ve sütun vektörleri

- $n \times 1$ matris bir n-vektördür
- 1×1 matris bir (skaler) sayıdır
- $1 \times n$ matris bir satır vektördür, örneğin

$$\begin{bmatrix} 1.2 & -0.3 & 1.4 & 2.6 \end{bmatrix}$$

ile verilen satır vektör

$$\begin{bmatrix} 1.2 \\ -0.3 \\ 1.4 \\ 2.6 \end{bmatrix}$$

ile verilen (sütun) vektör ile aynı değildir

Bir matrisin satır ve sütunları

- A'nın bir $m \times n$ matris olduğunu farz edelim (elemanları A_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$)
- A'nın j. sütunu
$$\begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix}$$
ile verilen m-vektördür
- A'nın i. satırı
$$\begin{bmatrix} A_{i1} & A_{i2} & \cdots & A_{in} \end{bmatrix}$$
ile verilen n-satır-vektördür

Matris dilimi

matris dilimi (*slice*): $A_{p:q,r:s}$

$$\begin{bmatrix} A_{pr} & A_{p,r+1} & \cdots & A_{ps} \\ A_{p+1,r} & A_{p+1,r+1} & \cdots & A_{p+1,s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{qr} & A_{q,r+1} & \cdots & A_{qs} \end{bmatrix}$$

ile verilen $(q-p+1) \times (s-r+1)$ matristir

Blok matrisler

- elemanları matrislerden oluşan blok matrisler kurabiliriz.

$$\text{örnek:} \quad A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}$$

  buradaki B, C, D ve E matrislerine A'nın blokları (veya altmatrisleri (*submatrix*)) denir
- her blok satırındaki matrislerin yüksekliği (yani, satır boyutu) aynı olmak zorundadır
- her blok sütunundaki matrislerin genişliği (yani, sütun boyutu) aynı olmak zorundadır
- örnek:

$$B = \begin{bmatrix} 0 & 2 & 3 \end{bmatrix}, \; C = \begin{bmatrix} -1 \end{bmatrix}, \; D = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 5 \end{bmatrix}, \; E = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix} = \begin{bmatrix} 0 & 2 & 3 & -1 \\ 2 & 2 & 1 & 4 \\ 1 & 3 & 5 & 4 \end{bmatrix}$$

Matrisin satır ve sütun gösterimi

- A bir $m \times n$ matris olsun
- A'yı bloklar m-vektör sütunları $a_1, a_2, \ldots, a_n$ olacak şekilde gösterebiliriz:

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

- veya, bloklar n-satır-vektör satırları $b_1, b_2, \ldots, b_m$ olacak şekilde gösterebiliriz:

$$A = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Örnekler

- görüntü: tek renkli (*monochrome*) görüntüde X_{ij} i, j'deki piksel değeri
- yağış verisi: A_{ij} i konumunda j tarihindeki yağış miktarı
- çoklu varlık getirileri: R_{ij} varlık j'nin dönem i'deki getirisi
- olumsallık (*contingency*) tablosu: A_{ij} birinci niteliği (*attribute*) i ve ikinci niteliği j olan cisimlerin sayısı
- öznitelik matrisi: X_{ij} öge (*entity*) j için öznitelik i'nin değeri

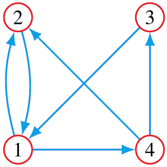
bu örneklerin her biri için, satır ve sütunların ne anlama geldiğini doğru anladığınızdan emin olun

Çizge ve bağıntı

- bağıntı (*relation*), $1, 2, \ldots, n$ ile etiketlenmiş cisim çiftlerinden oluşan bir kümedir, örneğin

$$\mathcal{R} = \{(1,2), (1,3), (2,1), (2,4), (3,4), (4,1)\}$$

- bağıntı, yönlü çizge (*directed graph*) ile aynıdır



- yönlü çizge, $n \times n$ matris olarak temsil edilebilir ($(i,j) \in \mathcal{R}$ ise $A_{ij} = 1$ ile)

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

Özel matrisler

- bütün elemanları 0 olan $m \times n$ matrise **sıfır matris** (*zero matrix*) denir. $0_{m \times n}$ veya kısaca 0 ile gösterilir
- $I_{ii} = 1$ ve $i \neq j$ için $I_{ij} = 0$ ile tanımlanan kare matrise **birim matris** (*identity matrix*) denir. I_n veya kısaca I ile gösterilir. örnek:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- çoğu elemanı sıfır olan matrise seyrek matris denir
  - örnekler: 0 ve I
  - verimli şekilde saklanabilir ve işlenebilirler
  - A isimli bir matrisin sıfır olmayan elemanlarının sayısı $\mathbf{nnz}(A)$ ile gösterilir

Köşegen ve üçgen matrisler

- köşegen (*diagonal*) matris: $i \neq j$ için $A_{ij} = 0$ ile tanımlanan kare matris
- $\mathbf{diag}(a_1, a_2, \ldots, a_n)$, $A_{ii} = a_i$ $(i = 1, 2, \ldots, n)$ ile tanımlanan köşegen matrisi ifade eder
- örnek:

$$\mathbf{diag}(0.2, -3, 1.2) = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1.2 \end{bmatrix}$$

- alt üçgen (*lower triangular*) matris: $i < j$ için $A_{ij} = 0$
- üst üçgen (*upper triangular*) matris: $i > j$ için $A_{ij} = 0$
- örnekler:

$$\begin{bmatrix} 1 & -1 & 0.7 \\ 0 & 1.2 & -1.1 \\ 0 & 0 & 3.2 \end{bmatrix} \text{ üst üçgen,} \quad \begin{bmatrix} -0.6 & 0 \\ -0.3 & 3.5 \end{bmatrix} \text{ alt üçgen}$$

- not: köşegen ve üçgen matrisler, kare matrisin özel halleridir

Matrisin devriği

- $m \times n$ bir A matrisinin devriği (*transpose*) A^T ile gösterilir ve

$$(A^T)_{ij} = A_{ji}, \quad i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m$$

- örnek:

$$\begin{bmatrix} 0 & 4 \\ 7 & 0 \\ 3 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 7 & 3 \\ 4 & 0 & 1 \end{bmatrix}$$

- devirme işlemi sütunları satırlara, satırları da sütunlara çevirir
- $(A^T)^T = A$

Toplama, çıkarma ve skaler çarpım

- (vektörlerde olduğu gibi), aynı boyutlu matrislerle toplama ve çıkarma işlemi yapabiliriz:

$$(A+B)_{ij} = A_{ij} + B_{ij}, \quad i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, m$$

örnek: $$\begin{bmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{bmatrix} + \begin{bmatrix} 3 & 4 & 5 \\ 6 & -7 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 8 \\ 2 & -2 & 2 \end{bmatrix}$$

- çıkarma benzer şekilde yapılır
- skaler çarpım:

$$(\alpha A)_{ij} = \alpha A_{ij}, \quad i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, m$$

- işlemlerin çoğu özelliği kolayca görülebilir, örneğin:

$$A + B = B + A, \quad \alpha(A+B) = \alpha A + \alpha B, \quad (A+B)^T = A^T + B^T$$

Matris normu

- $m \times n$ matris A için matris (Frobenius) normu

$$\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2 \right)^{1/2}$$

  olarak tanımlanır
- $n = 1$ için vektör normu ile aynıdır
- norm özelliklerini sağlar:

$$\begin{aligned}
&1) \|\alpha A\| = |\alpha| \|A\| \\
&2) \|A + B\| \leq \|A\| + \|B\| \\
&3) \|A\| \geq 0 \\
&4) \text{ancak } A = 0 \text{ ise } \|A\| = 0
\end{aligned}$$

- iki matrisin arasındaki uzaklık: $\|A - B\|$
- Frobenius normundan başka matris normları da vardır ancak biz bu derste onları kullanmayacağız

Bölüm 2

Matris-vektör çarpımı

Matris-vektör çarpımı

- $m \times n$ matris A ile n-vektör x'in matris-vektör çarpımı $y = Ax$ ile gösterilir ve

$$y_i = A_{i1}x_1 + A_{i2}x_2 + \cdots + A_{in}x_n, \quad i = 1, 2, \ldots, m$$

  olarak hesaplanır
- örnek:

$$\begin{bmatrix} 0 & 2 & -1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

Satır yorumu

- $y = Ax$ ifadesi

$$y_i = b_i^T x, \quad i = 1, 2, \ldots, m$$

  olarak gösterilebilir. burada $b_1^T, b_2^T, \ldots, b_m^T$ A'nın satırlarıdır
- dolayısıyla $y = Ax$ A'nın bütün satırlarıyla x'in "yığın" (*batch*) iç çarpımıdır
- örnek: $A\mathbf{1}$ A'nın satır toplamlarının vektörüdür

Sütun yorumu

- $y = Ax$ ifadesi

  $$y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

  olarak gösterilebilir. burada $a_1, a_2, \ldots, a_n$ A'nın sütunlarıdır
- dolayısıyla $y = Ax$ $x_1, x_2, \ldots, x_n$ katsayılarıyla (x'in elemanları) A'nın sütunlarının doğrusal bileşimidir
- önemli örnek: $Ae_j = a_j$
- $x = 0$ $Ax = 0$'ı gerektiriyorsa A'nın sütunları doğrusal bağımsızdır

Bölüm 3

Örnekler

Genel örnekler

- $0x = 0$: sıfır matrisiyle çarpım sıfır verir
- $Ix = x$: birim matrisle çarpım etkisizdir
- iç çarpım $a^T b$ $1 \times n$ matris a^T ile n-vektör b'nin matris-vektör çarpımıdır
- $\tilde{x} = Ax$ x'in ortalamadan arındırılmış halidir

$$A = \begin{bmatrix} 1 - 1/n & -1/n & \cdots & -1/n \\ -1/n & 1 - 1/n & \cdots & -1/n \\ \vdots & \vdots & \ddots & \vdots \\ -1/n & -1/n & \cdots & 1 - 1/n \end{bmatrix}$$

Fark matrisi

- $(n-1) \times n$ fark (*difference*) matrisi

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ & & \ddots & \ddots & & & \\ & & & \ddots & \ddots & & \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

- $y = Dx$ x'in ardışık elemanlarının farklarını içeren $(n-1)$-vektördür
- Dirichlet enerjisi $\|Dx\|^2$: bir zaman serisi x için dalgalılığın bir ölçüsüdür

Getiri matrisi - portföy vektörü

- $T \times n$ R varlık getirilerinin matrisi
- R_{ij} varlık j'nin dönem i'deki getirisi
- n-vektör w portföyü ifade eder (varlıklara yapılan yatırımlar)
- T-vektör Rw portföy getirisinin zaman serisi
- $\mathbf{avg}(Rw)$ portföy (ortalama) getirisi, $\mathbf{std}(Rw)$ portföy riski

Öznitelik matrisi - ağırlık vektörü

- $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ $n \times N$ öznitelik matrisi
- satır x_j cisim veya örnek j için öznitelik n-vektörü
- X_{ij} örnek j için öznitelik i'nin değeri
- n-vektör w ağırlık vektörü
- $s = X^T w$ her örnek için skorların vektörü; $s_j = x_j^T W$

Giriş-çıkış matrisi

- A $m \times n$ matris
- $y = Ax$
- n-vektör x *giriş* veya *etki*
- m-vektör y *çıkış* veya *sonuç*
- A_{ij} y_i'nin x_j'e nasıl bağlı olduğunu gösteren çarpan (*factor*)
- A_{ij} giriş j'den çıkış i'ye olan kazanç (*gain*)
- örneğin, A alt üçgen ise y_i sadece $x_1, x_2, \ldots, x_i$'ye bağlıdır

Karmaşıklık

- $m \times n$ matris A $m \times n$ sayı dizilimi olarak saklanır (seyrek A için sadece $\mathbf{nnz}(A)$ adet sıfır olmayan değer saklanır)
- matris toplamanın ve skaler-matris çarpımın maliyeti mn flop
- matris-vektör çarpımının maliyeti $m(2n-1) \approx 2mn$ flop (seyrek A için yaklaşık $2\mathbf{nnz}(A)$ flop)