Doğrusal Cebir ve Kalkülüs hatırlatması VIP

Afshine Amidi ve Shervine Amidi

April 30, 2019

Kadir Tekeli ve Ekrem Çetinkaya tarafından çevrilmiştir

Genel notasyonlar

❒ **Vektör** – i-inci elemanı $x_i \in \mathbb{R}$ olmak üzere n elemanlı bir vektör, $x \in \mathbb{R}^n$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

❒ **Matris** – $A_{i,j} \in \mathbb{R}$ i-inci satır ve j-inci sütundaki elemanları olmak üzere m satırlı ve n sütunlu bir matris, $A \in \mathbb{R}^{m\times n}$:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

Dipnot: Yukarıda tanımlanan x vektörü $n\times 1$ tipinde bir matris olarak ele alınabilir ve genellikle sütun vektörü olarak adlandırılır.

❒ **Birim matris** – Birim matris, köşegeni birlerden ve diğer tüm elemanları sıfırlardan oluşan karesel matris, $I \in \mathbb{R}^{n\times n}$:

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Dipnot: Her $A \in \mathbb{R}^{n\times n}$ matrisi için $A \times I = I \times A = A$ eşitliği sağlanır.

❒ **Köşegen matris** – Bir köşegen matris, köşegenindeki elemanları sıfırdan farklı diğer tüm elemanları sıfır olan karesel matris, $D \in \mathbb{R}^{n\times n}$:

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

Dipnot: D matrisi $\textrm{diag}(d_1,...,d_n)$ olarak da gösterilir.

Matris işlemleri

❒ **Vektör-vektör** – İki çeşit vektör-vektör çarpımı vardır.

- iç çarpım: $x,y \in \mathbb{R}^n$ için:

$$\boxed{x^Ty = \sum_{i=1}^n x_iy_i \in \mathbb{R}}$$

- dış çarpım: $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ için:

$$\boxed{xy^T = \begin{pmatrix} x_1y_1 & \cdots & x_1y_n \\ \vdots & & \vdots \\ x_my_1 & \cdots & x_my_n \end{pmatrix} \in \mathbb{R}^{m\times n}}$$

❒ **Matris-vektör** – $A \in \mathbb{R}^{m\times n}$ matrisi ve $x \in \mathbb{R}^n$ vektörünün çarpımları $\mathbb{R}^m$ boyutunda bir vektördür:

$$\boxed{Ax = \begin{pmatrix} a_{r,1}^Tx \\ \vdots \\ a_{r,m}^Tx \end{pmatrix} = \sum_{i=1}^n a_{c,i}x_i \in \mathbb{R}^m}$$

burada $a_{r,i}^T$ A'nın vektör satırları ve $a_{c,j}$ A'nın vektör sütunları ve x_i x vektörünün elemanlarıdır.

❒ **Matris-matris** – $A \in \mathbb{R}^{m\times n}$ matrisi ve $B \in \mathbb{R}^{n\times p}$ matrisinin çarpımları $\mathbb{R}^{n\times p}$ boyutunda bir matristir:

$$\boxed{AB = \begin{pmatrix} a_{r,1}^Tb_{c,1} & \cdots & a_{r,1}^Tb_{c,p} \\ \vdots & & \vdots \\ a_{r,m}^Tb_{c,1} & \cdots & a_{r,m}^Tb_{c,p} \end{pmatrix} = \sum_{i=1}^n a_{c,i}b_{r,i}^T \in \mathbb{R}^{n\times p}}$$

burada $a_{r,i}^T, b_{r,i}^T$ sırasıyla A ve B'nin vektör satırları ve $a_{c,j}, b_{c,j}$ sırasıyla A ve B'nin vektör sütunlarıdır.

❒ **Devrik (Transpoze)** – Bir $A \in \mathbb{R}^{m\times n}$ matrisinin devriği, satır ve sütunların yer değiştirmesi ile elde edilir, ve A^T ile gösterilir:

$$\boxed{\forall i,j, \quad A_{i,j}^T = A_{j,i}}$$

Dipnot: Her A,B için $(AB)^T = B^TA^T$ vardır.

❒ **Ters** – Tersinir bir A karesel matrisinin tersi, aşağıdaki koşulu sağlayan matristir, ve A^{-1} ile gösterilir:

$$\boxed{AA^{-1} = A^{-1}A = I}$$

Dipnot: Her karesel matris tersinir değildir. Ayrıca, Her tersinir A,B matrisi için $(AB)^{-1} = B^{-1}A^{-1}$ dir.

❒ **İz** – Bir A karesel matrisinin izi, köşegenindeki elemanlarının toplamıdır, ve $\textrm{tr}(A)$ ile gösterilir:

$$\boxed{\mathrm{tr}(A)=\sum_{i=1}^n A_{i,i}}$$

Dipnot: A,B matrisleri için $tr(A^T)=tr(A)$ ve $tr(AB)=tr(BA)$ vardır.

❒ **Determinant** – $A\in\mathbb{R}^{n\times n}$ matrisinin determinantı, $A_{\backslash i,\backslash j}$ gösterimi i-inci satırsız ve j-inci sütunsuz şekilde A matrisi olmak üzere özyinelemeli olarak aşağıdaki gibi ifade edilir, ve $|A|$ ya da $\det(A)$ ile gösterilir:

$$\boxed{\det(A)=|A|=\sum_{j=1}^n(-1)^{i+j}A_{i,j}|A_{\backslash i,\backslash j}|}$$

Dipnot: A tersinirdir ancak ve ancak $|A|\neq0$. Ayrıca, $|AB|=|A||B|$ ve $|A^T|=|A|$.

Matris özellikleri

❒ **Simetrik ayrışım** – Verilen bir A matrisi simetrik ve ters simetrik parçalarının cinsinden aşağıdaki gibi ifade edilebilir:

$$\boxed{A=\underbrace{\frac{A+A^T}{2}}_{\text{Simetrik}}+\underbrace{\frac{A-A^T}{2}}_{\text{Ters simetrik}}}$$

❒ **Norm** – V vektör uzayı ve her $x,y\in V$ için aşağıdaki özellikleri sağlayan $N:V\longrightarrow[0,+\infty[$ fonksiyonu bir normdur:

- $N(x+y)\leqslant N(x)+N(y)$
- Bir a sabiti için $N(ax)=|a|N(x)$
- $N(x)=0$ ise $x=0$

$x\in V$ için en yaygın şekilde kullanılan normlar aşağıdaki tabloda verilmektedir.

Norm	Notation	Definition	Use case
Manhattan, L^1	$\|\|x\|\|_1$	$\displaystyle\sum_{i=1}^n\|x_i\|$	LASSO regularization
Euclidean, L^2	$\|\|x\|\|_2$	$\displaystyle\sqrt{\sum_{i=1}^n x_i^2}$	Ridge regularization
p-norm, L^p	$\|\|x\|\|_p$	$\displaystyle\left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}$	Hölder inequality
Infinity, L^∞	$\|\|x\|\|_\infty$	$\underset{i}{\max}\ \|x_i\|$	Uniform convergence

❒ **Doğrusal bağımlılık** – Bir vektör kümesinden bir vektör diğer vektörlerin doğrusal birleşimi (kombinasyonu) cinsinden yazılabiliyorsa bu vektör kümesine doğrusal bağımlı denir.
Dipnot: Eğer bu şekilde yazılabilen herhangi bir vektör yoksa bu vektörlere doğrusal bağımsız denir.

❒ **Matris rankı** – Verilen bir A matrisinin rankı, $\mathrm{rank}(A)$, bu matrisinin sütunları tarafından üretilen vektör uzayının boyutudur. Bu ifade A matrisinin doğrusal bağımsız sütunlarının maksimum sayısına denktir.

❒ **Pozitif yarı-tanımlı matris** – Aşağıdaki koşulu sağlayan bir $A\in\mathbb{R}^{n\times n}$ matrisi pozitif yarı-tanımlıdır ve $A\succeq0$ ile gösterilir:

$$\boxed{A=A^T}\quad\text{and}\quad\boxed{\forall x\in\mathbb{R}^n,\quad x^TAx\geqslant0}$$

Dipnot: Benzer olarak, pozitif yarı-tanımlı bir A matrisi sıfırdan farklı her x vektörü için $x^TAx>0$ koşulunu sağlıyorsa A matrisine pozitif tanımlı denir ve $A\succ0$ ile gösterilir.

❒ **Özdeğer, özvektör** – Verilen bir $A\in\mathbb{R}^{n\times n}$ için aşağıdaki gibi bir $z\in\mathbb{R}^n\backslash\{0\}$ vektörü var ise buna özvektör, λ sayısına da A matrisinin öz değeri denir.

$$\boxed{Az=\lambda z}$$

❒ **Spektral teorem** – $A\in\mathbb{R}^{n\times n}$ olsun. Eğer A simetrik ise, A matrisi gerçel ortogonal $U\in\mathbb{R}^{n\times n}$ matrisi ile köşegenleştirilebilir. $\Lambda=\mathrm{diag}(\lambda_1,...,\lambda_n)$ olmak üzere:

$$\boxed{\exists\Lambda\text{ köşegen},\quad A=U\Lambda U^T}$$

❒ **Tekil-değer ayrışımı** – $m\times n$ tipindeki bir A matrisi için tekil-değer ayrışımı; $m\times m$ tipinde bir üniter U, $m\times n$ tipinde bir köşegen Σ ve $n\times n$ tipinde bir üniter V matrislerinin varlığını garanti eden bir parçalama tekniğidir.

$$\boxed{A=U\Sigma V^T}$$

Matris kalkülüsü

❒ **Gradyan** – $f:\mathbb{R}^{m\times n}\rightarrow\mathbb{R}$ bir fonksiyon ve $A\in\mathbb{R}^{m\times n}$ bir matris olsun. f nin A ya göre gradyanı $m\times n$ tipinde bir matristir, ve $\nabla_Af(A)$ ile gösterilir:

$$\boxed{\Big(\nabla_Af(A)\Big)_{i,j}=\frac{\partial f(A)}{\partial A_{i,j}}}$$

Dipnot: f fonksiyonunun gradyanı yalnızca f skaler döndüren bir fonksiyon ise tanımlıdır.

❒ **Hessian** – $f:\mathbb{R}^n\rightarrow\mathbb{R}$ bir fonksiyon ve $x\in\mathbb{R}^n$ bir vektör olsun. f fonksiyonun x vektörüne göre Hessian'ı $n\times n$ tipinde bir simetrik matristir, ve $\nabla_x^2f(x)$ ile gösterilir:

$$\boxed{\Big(\nabla_x^2f(x)\Big)_{i,j}=\frac{\partial^2f(x)}{\partial x_i\partial x_j}}$$

Dipnot: f fonksiyonunun Hessian'ı yalnızca f skaler döndüren bir fonksiyon ise tanımlıdır.

❒ **Gradyan işlemleri** – A,B,C matrisleri için aşağıdaki işlemlerin akılda bulunmasında fayda vardır:

$$\boxed{\nabla_A\mathrm{tr}(AB)=B^T}\qquad\boxed{\nabla_{A^T}f(A)=(\nabla_Af(A))^T}$$

$$\boxed{\nabla_A\mathrm{tr}(ABA^TC)=CAB+C^TAB^T}\qquad\boxed{\nabla_A|A|=|A|(A^{-1})^T}$$