Hesaplamalı Tarifler II: Matris Ayrışımı

Ş. İlker Birbil / sibirbil@sabanciuniv.edu.tr / www.bolbilim.com

Yaklaşık 2.200 yıldır denklem sistemlerini çözmeye çalışıyoruz [1]. Hemen öyle karmaşık sistemler gelmesin aklınıza. Doğrusal eşitliklerden oluşan, görece kolay sistemlerin çözülmeleri bile hâlâ büyük bir olay. Çünkü mühendislik ve doğa bilimleri başta olmak üzere, hesaplamalı işlerin tamamında daima karşımıza bir denklem sistemi çıkıyor. Sadece her geçen yıl daha büyük sistemleri çözüyoruz. Fakat ne kadar devasa olurlarsa olsunlar, neticede bu sistemler dönüp dolaşıp bizim oğlanın eve getirdiği şu ödeve benziyor:

> "Ayşe, annesi ve babasının yaşları toplamı 80'dir. Ayşe'nin ve babasının yaşlarının 2 katları ile annesinin yaşını toplayınca 120 ediyor. Ayşe'nin yaşının 3 katı ile babasının yaşının 2 katının toplamından, annesinin yaşını çıkarınca geriye 48 kalıyor. Ayşe'nin, annesinin ve babasının yaşları kaçtır?"

Hah, çok kolay. Matematik Dünyası okuyucusu basit cebirden mi korkacak? Ayşe'nin, babasının ve annesinin yaşları sırasıyla, x_1, x_2 ve x_3 olsun. Altı üstü üç bilinmeyenli üç denklemi çözeceğiz:

$$\begin{aligned} x_1 + x_2 + x_3 &= 80, \\ 2x_1 + 2x_2 + x_3 &= 120, \\ 3x_1 + 2x_2 - x_3 &= 48. \end{aligned}$$

Matris cebiriyle ifade edersek

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 2 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 80 \\ 120 \\ 48 \end{bmatrix}}_{d}.$$

Yani

$$Ax = d \tag{1}$$

ile gösterilen ve çözülmeyi bekleyen bir sistem var karşımızda. Ya sonra? Liseden hatırlarsınız belki. Satır işlemleri ile bilinmeyenlerden birini çekmek için çeşitli numaralar yapılabilir. Biraz uzun iş. Oysa birisi bize A matrisini şöyle çarpanlarına ayırıp verse ne güzel olur:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 2 & -1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \frac{1}{3} & \frac{1}{2} & 1 \\ \frac{2}{3} & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} 3 & 2 & -1 \\ 0 & \frac{2}{3} & \frac{5}{3} \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{C}.$$

Güzel olur çünkü hayatımızı çok kolaylaştırır. O sayede çözeceğimiz (1) sistemini

$$Ax = B\underbrace{Cx}_{y} = d$$

şeklinde baştan yazabiliriz. Bir kenara $Cx = y$ diye not alıp, $By = d$ sistemine bir bakalım:

$$\underbrace{\begin{bmatrix} \frac{1}{3} & \frac{1}{2} & 1 \\ \frac{2}{3} & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} 80 \\ 120 \\ 48 \end{bmatrix}}_{d}.$$

B matrisinin köşegeni altındaki değerler 0 olduğu için sırayla değerleri yerine koyarak sistemi çözebiliriz. En son satırdan başlarsak

$$\begin{aligned} & y_1 = 48, \\ & \tfrac{2}{3}y_1 + y_2 = 120 \implies y_2 = 120 - 32 = 88, \\ & \tfrac{1}{3}y_1 + \tfrac{1}{2}y_2 + y_3 = 80 \implies y_3 = 80 - 16 - 44 = 20, \end{aligned}$$

işlemleriyle y vektörünü bulmuş oluruz. Bir kenara $Cx = y$ yazmıştık. O zaman

$$\underbrace{\begin{bmatrix} 3 & 2 & -1 \\ 0 & \frac{2}{3} & \frac{5}{3} \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 48 \\ 88 \\ 20 \end{bmatrix}}_{y}$$

eşitliğini elde ederiz. İşin güzel tarafı bu sefer de C matrisinin köşegeni altındaki değerler 0. Tıpkı önceki gibi değerleri yerine koyarak ilerleyebiliriz. Ayşe'nin annesinin yaşını son satırdan buluruz. Zaten diğerleri de çorap söküğü gibi gelir:

$$\begin{aligned} & \tfrac{1}{2}x_3 = 20 \implies x_3 = 40, \\ & \tfrac{2}{3}x_2 + \tfrac{5}{3}x_3 = 88 \implies x_2 = \tfrac{3}{2}(88 - \tfrac{200}{3}) = 32, \\ & 3x_1 + 2x_2 - x_3 = 48 \implies x_1 = \tfrac{1}{3}(48 - 64 + 40) = 8. \end{aligned}$$

Açıkçası $A = BC$ ifadesini sağlayan B ve C matrisleri verilince problemi çözmek çocuk oyuncağı oldu. İyi de B ve C matrislerini nasıl elde edeceğiz? Bu sorunun cevabı, bizi bu sayıdaki hesaplamalı tarife getiriyor. Bir matrisi çarpanlarına ayırmak için önerilmiş farklı yöntemler var. Aralarından iki tanesini bu yazıda konuşacağız: LU ve Cholesky ayrışımı. Bu yöntemleri seçmemin iki sebebi var. Bir, ikisi de oldukça sık kullanılan yöntemler. İki, matris ayrışımı konusuna bir giriş yapmak için son derece uygunlar. Tarifi anlatmaya başlamadan önce malzemelerimizi bir listeleyelim.

1 Malzemeler

Konuya giriş yaparken matrisleri kullandık bile. Birkaç sembolü daha tanıtırsak ilerleyen bölümleri yazmamız kolaylaşır. Diyelim elimizde m satırdan ve n sütundan oluşan bir A matrisi var. Boyutlarını görmek için $A_{m\times n}$ şeklinde yazabiliriz. Tüm elemanları 0 olan bir matrisi $0_{m\times n}$ ile gösterebiliriz. Eğer bir matrisin (i,j) elemanından bahsediyorsak, bu i. satır ve j. sütundaki değer demektir. Matrisler için çoğunlukla büyük harfler kullanılırken, elemanları küçük harfle gösterilir. Örneğin A matrisini yazdığımızda a_{ij} ile de matrisin (i,j) elamanını kastederiz:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Matris elemanlarının yerlerini belirlediğimize göre devrik matrisi tanımlayabiliriz. Yukarıda verdiğimiz A matrisinin devriği $A^\intercal$ ile gösterilir. Bu durumda (i,j) elemanı, (j,i) elemanı ile yer değiştirir. Kısacası A matrisinin satırları sütun, sütunları satır olunca $A^\intercal$ matrisi elde edilir. Yani

$$A^\intercal = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

olur. Kare matrislerin –adı üstünde– satır sayısı sütun sayısına eşittir. Bir kare matrisin simetrik olması devriğinin kendisine eşit olması demektir. Yeni tanıştığımız sembollerle gösterirsek, $A_{n\times n}$ matrisi simetrik ise $A = A^\intercal$ eşitliğini sağlar.

Köşegen bir matrisin sadece $(1,1)$ ile başlayıp (m,n) ile biten köşegenindeki sayıları bilmek yeterlidir. Çünkü köşegenleri haricindeki tüm elemanlar 0 olur:

$$D = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & d_{mn} \end{bmatrix}.$$

Köşegen matrislerin en meşhuru birim matristir. Birim matris, köşegeni üzerindeki tüm elemanlar 1 değerini alan kare bir matristir. Boyutunu da belirtmek için kısaca I_n ile gösterilir.

Sıra geldi matrislerde satır ve sütun işlemlerine. Satırlar ile başlayalım. Diyelim bir matrisin ilk satırını $\frac{2}{3}$ ile çarpıp ikinci satırdan çıkaracağız. Bir örnek üzerinde hemen gösterelim:

$$\begin{bmatrix} 9 & 12 & 24 \\ 6 & 6 & 8 \\ 3 & 4 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix}.$$

İşin güzel tarafı bu işlemin aynısını birim matris üzerinde yapıp, elde ettiğimiz matrisi, asıl matrisimizle soldan çarpabiliriz. Bu da aynı sonucu verecektir:

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 9 & 12 & 24 \\ 6 & 6 & 8 \\ 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix}.$$

Birim matristen satır işlemleriyle elde ettiğimiz ve soldan çarparken kullandığımız bu matrislere yalın matrisler denir. Eğer bir matrisin satırlarını yer değiştirmek istersek, benzer şekilde aynı işlemi birim matrisin satırları üzerinde yapıp yine soldan asıl matrisimizle çarpabiliriz. Birim matristen türettiğimiz bu matrislere ise permütasyon matrisleri denir. Sütun işlemlerine gelirsek, bu işlemler satır işlemlerine çok benzer. Bu sefer tek yapacağımız bir matrisi soldan değil de, sağdan yalın matrislerle ya da permütasyon matrisleri ile çarpmak olur. O durumda hem yalın matrisler, hem de permütasyon matrisleri birim matrisin sütunları üzerinde ilgili işlemlerin yapılması ile elde edilirler.

Bir kare matrisin sadece köşegeni ve altındaki değerler 0'dan farklıysa o matrise alt üçgen matris denir. Böyle bir matrisin devriği de üst üçgen matris olur. Eğer alt üçgen matrisleri birbirleriyle çarparsak, ortaya yine alt üçgen bir matris çıkar. Ufak bir örnekte kolayca deneyebilirsiniz.

Bazı durumlarda bir matrisi alt matrislere bölerek blok matrisler ile çalışmak isteyebiliriz. Örneğin $A_{m\times n}$ matrisini şu şekilde 4 alt matrise bölebiliriz:

$$A = \begin{bmatrix} E_{m_1\times n_1} & F_{m_1\times(n-n_1)} \\ G_{(m-m_1)\times n_1} & H_{(m-m_1)\times(n\times n_1)} \end{bmatrix}.$$

Gelelim bir kare matrisin tersine. Verilen $A_{n\times n}$ matrisi için

$$AA^{-1} = A^{-1}A = I_n$$

eşitliğini sağlayan bir A^{-1} matrisi bulabiliyorsak, A matrisinin tersi var demektir. Burada A^{-1} sembolü ile A matrisinin tersi gösterilir. Bir matrisin –eğer varsa– tersi tektir. Tersi olmayan matrislere ise tekil matrisler denir. Bir matris tekil değilse, çarpanları da tekil değildir. Tekil olmayan bir matrisi herhangi bir permütasyon matrisi çarparak elde ettiğimiz matrisin de tersi mevcuttur. Tersi olan bir matrisin, bir sütunu ya da bir satırı tamamen sıfır olamaz. Yalın alt üçgen matrislerin tersleri de yalın alt üçgen olurlar. Bu son dört cümlenin kanıtlanması güzel olur. Bir deneyin bakalım. Blok matrislere dönersek, elimizde

$$A = \begin{bmatrix} E_{k\times k} & F_{k\times(n-k)} \\ 0_{(n-k)\times k} & H_{(n-k)\times(n\times k)} \end{bmatrix}.$$

şeklinde verilmiş bir üst üçgen matris olsun. Bu matrisin tersinin olması için gerek ve yeter şart, köşegenleri üzerindeki E ve H matrislerinin tekil olmamalarıdır [1].

Pozitif belirli matrisleri de tanıtırsak malzemelerimiz tamamlanmış olacak. Elimizde $A_{n\times n}$ olarak verilmiş bir kare matris olsun. Eğer her $y \neq 0$ vektörü için

$$y^{\mathsf{T}} A y > 0$$

eşitsizliği sağlanıyorsa, A matrisi pozitif belirli demektir. Pozitif belirli bir matrisin köşegenindeki değerlerin tamamı pozitiftir. Daha da güzeli, pozitif belirli bir matrisin tersi her zaman vardır. Bir matrisin pozitif belirli olduğunu bulmanın başka yolları da mevcut. Ancak bu yazının devamı için bu tanım kâfi gelir. Bir diğer tanımı zaten Bölüm 3'te vereceğiz.

2 LU Ayrışımı

İlk konuşacağımız ayrışımın amacı, tersi olduğu bilinen bir $A_{n\times n}$ matrisini, alt üçgen bir $L_{n\times n}$ matrisi ile üst üçgen bir $U_{n\times n}$ matrisinin çarpımı şeklinde yazmak:

$$A = LU.$$

Bu sonucu elde etmek için önce A matrisinin köşegeni altındaki değerleri 0'a eşitlemeye çalışacağız. Bunun için de sırasıyla birkaç satır işlemi yapacağız. Temel fikri vermek için, bu adımlar sırasında satırların yerlerinin değişmediğini varsayalım. Kaldı ki yerlerinin değiştiği durum da benzer şekilde yapılabilir. O konuya bu bölümün sonundaki teorem ile geleceğim.

Yine bir örnek ile başlamakta fayda var. Diyelim ki matrisimiz

$$A = \begin{bmatrix} 9 & 12 & 24 \\ \boxed{6} & 6 & 8 \\ \boxed{3} & \boxed{4} & 4 \end{bmatrix}$$

şeklinde verilmiş olsun. Matriste kutu içine aldığım, köşegen altındaki değerleri 0 yapmaya çalışacağız. Bu sayede üst üçgen U matrisini elde edebiliriz. İlk satırı $\frac{2}{3}$ ile çarpıp ikinci satırdan çıkaralım:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_1}\underbrace{\begin{bmatrix} 9 & 12 & 24 \\ 6 & 6 & 8 \\ 3 & 4 & 4 \end{bmatrix}}_{A} = \begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix}.$$

Ardından yine ilk satırı alıp, $\frac{1}{3}$ ile çarpıp son satırdan çıkararak

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{3} & 0 & 1 \end{bmatrix}}_{E_2}\underbrace{\begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix}}_{E_1 A} = \underbrace{\begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 0 & 0 & -4 \end{bmatrix}}_{U}$$

eşitliğini buluruz. Şansımıza $(3,2)$ konumunda da 0 elde ettik. Eğer 0 olmasaydı, ikinci satırı kullanarak o konumda da 0 çıkmasını sağlardık. Bu örnek için ikinci satırı 0 ile çarpıp, son satırdan çıkardık diye düşünebiliriz. Son durumda elimizde şu sistem var:

$$E_2 E_1 A = U.$$

Dikkat ederseniz, satır değişikliği yapmadığımız için yalın matrislerimizin tamamı alt üçgen. Bir önceki bölümde bahsettiğimiz gibi bu matrislerin tersleri de alt üçgen olur. O zaman istediğimiz sonucu elde ettik demektir:

$$A = \underbrace{E_1^{-1} E_2^{-1}}_{L} U.$$

Burada L matrisini açıkça yazacak olursak

$$L = E_1^{-1} E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \boxed{\frac{2}{3}} & 1 & 0 \\ \boxed{\frac{1}{3}} & \boxed{0} & 1 \end{bmatrix}$$

eşitliğini elde ederiz. Bir noktaya dikkat etmenizi istiyorum. A matrisinden U matrisine ulaşmaya çalışırken, satırları tam da L matrisinde kutu içersinde gösterdiğim sayılar ile çarpmıştık. Yani A matrisinin ikinci satırında kutu içindeki 6 değerini 0 yapmak için, A matrisinin ilk satırını L matrisinde aynı konumdaki $\frac{2}{3}$ ile çarptık. Benzer şekilde A matrisindeki 3 değerini 0 yapmak için L matrisinde aynı konumdaki $\frac{1}{3}$ çarpanını kullandık. Yapmamıza gerek kalmadı ama A matrisinin son satırında kutu içindeki 4 değerini sıfırlamak için L matrisinde aynı konumdaki 0 değerini kullandık diyebiliriz. Ayrıca üçgen yalın matrisleri birbirleriyle çarptığımız için L matrisinin köşegeni sadece 1 değerlerinden oluştu.

Bu gözlem önemli. Çünkü başka herhangi bir matris kullanmadan, doğrudan A matrisinin üzerine L ve U matrislerinin değerlerini yazabiliriz. Deneyelim:

(i) İlk satırı $\frac{2}{3}$ ile çarp ve ikinci satırdan çıkar:

$$\underbrace{\begin{bmatrix} 9 & 12 & 24 \\ 6 & 6 & 8 \\ 3 & 4 & 4 \end{bmatrix}}_{A} \rightarrow \begin{bmatrix} 9 & 12 & 24 \\ \boxed{\frac{2}{3}} & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix}.$$

(ii) İlk satırı $\frac{1}{3}$ ile çarp ve son satırdan çıkar:

$$\begin{bmatrix} 9 & 12 & 24 \\ \boxed{\frac{2}{3}} & -2 & -8 \\ 3 & 4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 9 & 12 & 24 \\ \boxed{\frac{2}{3}} & -2 & -8 \\ \boxed{\frac{1}{3}} & 0 & -4 \end{bmatrix}.$$

(iii) Bu adımı yapmaya gerek kalmıyor ama yine de yazalım. İkinci satırı 0 ile çarp ve son satırdan çıkar:

$$\begin{bmatrix} 9 & 12 & 24 \\ \frac{2}{3} & -2 & -8 \\ \frac{1}{3} & 0 & -4 \end{bmatrix} \rightarrow \begin{bmatrix} 9 & 12 & 24 \\ \boxed{\frac{2}{3}} & -2 & -8 \\ \boxed{\frac{1}{3}} & \boxed{0} & -4 \end{bmatrix}.$$

En son elde ettiğimiz bu matristeki kutu içine yazılmış sayılar alt üçgen L matrisinin köşegeni altındaki değerleri, kalanlar da üst üçgen U matrisinin 0'dan farklı değerlerini veriyor.

Bir örnek verip de işin içinden sıyrılmak olmaz. LU ayrışımının her zaman çalışacağını matematiksel olarak da göstermeliyiz. Takip eden teorem tam da bunu yapıyor. Bu teoremin kare olmayan matrisler için daha genel hali yazılabilir. Zaten ben de öyle bir teoremden yararlandım [2].

Teorem 1. *Elimizde tekil olmayan bir $A_{n\times n}$ matrisi olsun. Bu matrisi, $P_{n\times n}$ ile gösterilen bir permütasyon matrisinin yardımıyla*

$$PA = LU \qquad (2)$$

olarak yazabiliriz. Burada L köşegen değerleri bir olan alt üçgen bir matris, U ise köşegen değerleri 0'dan farklı bir üst üçgen matristir. Ayrıca P matrisi seçildikten sonra (2) eşitliğini sağlayacak L ve U matris çifti tektir.

Kanıt: Matrisin boyutları üzerinden tümevarım yöntemini kullanabiliriz. Başlangıç için $n = 1$ alalım. Bu durumda $L = P = 1$ ve $U = A \neq 0$ olur. Şimdi $n > 1$ olduğu durumu inceleyelim. Elimizdeki A matrisi tekil olmadığı için ilk sütununda 0 olmayan en az bir eleman vardır. B durumda $\hat{P}$ matrisini öyle bir seçebiliriz ki $\hat{P}A$ matrisinin $(1,1)$ elemanı 0 olmaz. Kısacası, matrisi alt matrisler kullanarak yazabiliriz:

$$\hat{P}A = \begin{bmatrix} a & D^{\intercal} \\ C & B \end{bmatrix}.$$

Burada $a \neq 0$, $C_{(n-1)\times 1}$ ve $D_{(n-1)\times 1}$ ve $B_{(n-1)\times(n-1)}$ olur. İki çarpan matrisi yardımıyla şu eşitliği yazabiliriz:

$$\begin{bmatrix} a & D^{\intercal} \\ C & B \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a^{-1}C & I_{n-1} \end{bmatrix} \begin{bmatrix} a & D^{\intercal} \\ 0 & B - a^{-1}CD^{\intercal} \end{bmatrix}.$$

Bu eşitliğin solundaki matris tekil olmadığı için çarpanları da tekil değildir. Bu durumda eşitliğin sağındaki üst üçgen matrisin alt köşegenindeki $(n-1)\times(n-1)$ boyutlu matris de tekil olamaz. O halde, tümevarım hipotezimize göre $\bar{P}$ permütasyon matrisi kullanılarak elde edilen bir LU ayrışımı mevcuttur. Yani

$$\bar{P}(B - a^{-1}cd^{\intercal}) = \bar{L}\bar{U}$$

olur. Bu durumda

$$P = \begin{bmatrix} 1 & 0_{1\times(n-1)} \\ 0_{(n-1)\times 1} & \bar{P} \end{bmatrix} \hat{P}$$

matrisi ile teoreme uygun

$$L = \begin{bmatrix} 1 & 0_{1\times(n-1)} \\ a^{-1}\bar{P}C & \bar{L} \end{bmatrix}$$

ve

$$U = \begin{bmatrix} a & D^{\intercal} \\ 0_{(n-1)\times 1} & \bar{U} \end{bmatrix}$$

üçgen matrislerini kullanarak (2) eşitliğini elde ederiz.

Kanıtımızın son kısmında (2) ile gösterilen denklemdeki L ve U matris çiftinin tek olduklarını gösterelim. Varsayalım ki A matrisinin $PA = LU$ ve $PA = \bar{L}\bar{U}$ şeklinde iki LU ayrışımı olsun. Bu durumda $LU = \bar{L}\bar{U}$ olur. A matrisi tekil olmadığına göre, L ve U çarpanları da tekil değildir. Dolayısıyla

$$\bar{L}^{-1}L = \bar{U}U^{-1}$$

eşitliğini yazabiliriz. Bu eşitliğin bir tarafında alt üçgen, diğer tarafında ise üst üçgen bir matris var. O nedenle eşitlik sadece çarpımların köşegen matris verdikleri durumda sağlanır. L matrisinin köşegen değerleri 1 olduğuna göre

$$\bar{L}^{-1}L = \bar{U}U^{-1} = I_n$$

eşitliğini ve dolayısıyla $L = \bar{L}$ ve $U = \bar{U}$ sonucunu göstermiş oluruz. □

Burada önemli bir noktanın altını çizmeliyim. Teoremde verilen (2) denklemine göre, A matrisi üzerinde satır yerlerini değiştirme işlemleri de yapıyoruz. Kabaca söylemek gerekirse bu işlemlerle, çalışılan matrisin sol üst köşesinde 0'dan farklı bir değer elde ediyoruz.

3 Cholesky Ayrışımı

LU ayrışımı iyi hoş ama ufak bir garipliği var. L matrisinin köşegen değerleri 1 iken U matrisinin köşegen değerleri 1 değil. Simetrik bir çözüm elde etsek güzel olur. Bunun için U matrisinin her satırını, o satıra karşılık gelen köşegen değere bölebiliriz. Bölmede kullandığımız sayıları da bir köşegen matrise taşırız ve yeni U matrisi ile soldan çarparız. Önceki örneğimize dönersek

$$\begin{bmatrix} 9 & 12 & 24 \\ 0 & -2 & -8 \\ 0 & 0 & -4 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1 & \frac{12}{9} & \frac{24}{9} \\ 0 & 1 & \frac{-8}{-2} \\ 0 & 0 & 1 \end{bmatrix}$$

eşitliğini elde ederiz. Bu durumda A matrisi, üç matrisin çarpımı biçiminde yazılabilir:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 0 \\ \frac{1}{3} & 0 & 1 \end{bmatrix} \begin{bmatrix} 9 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1 & \frac{12}{9} & \frac{24}{9} \\ 0 & 1 & \frac{-8}{-2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Genelleyecek olursak, önceki bölümde anlattığımız gibi $A = LU$ şeklinde bir LU ayrışımı verilmiş olsun. Bu durumda

$$D = \begin{bmatrix} u_{11} & 0 & \dots & 0 \\ 0 & u_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

ve

$$\hat{U} = \begin{bmatrix} 1 & \frac{u_{12}}{u_{11}} & \frac{u_{13}}{u_{11}} & \dots & \frac{u_{1n}}{u_{11}} \\ 0 & 1 & \frac{u_{23}}{u_{22}} & \dots & \frac{u_{2n}}{u_{22}} \\ 0 & 0 & 1 & \dots & \frac{u_{3n}}{u_{33}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

matrislerini kullanarak

$$A = LD\hat{U}$$

eşitliğini yazabiliriz. Eğer A matrisinin simetrik olduğu da biliniyorsa

$$A = LDL^\intercal \tag{3}$$

ifadesi elde edilir. Bu son ifadeyi kanıtlamayacağım. Güzel bir ev ödevi olur ama.

Artık bu bölümün konusu olan ayrışım yöntemine geçebiliriz. Derdimiz matrislerin köklerini almak. "Kök almak" ifadesi yeterince açık olmadı. Diyelim elimizde simetrik ve pozitif belirli bir A matrisi var. Bu matrisi

$$A = R^\intercal R \tag{4}$$

şeklinde yazmaya çalışacağız. Bu işleme Cholesky ayrışımı adı verilir. Dikkate ederseniz R matrisi, A matrisinin karekökü gibi duruyor.

Sıradaki teorem bir taşla iki kuş vuruyor. Hem Cholesky ayrışımın nasıl yapılacağını gösteriyor, hem de bu ayrışımın pozitif belirlilik testi için biçilmiş kaftan olduğunu söylüyor.

Teorem 2. *Simetrik bir $A_{n\times n}$ matrisinin pozitif belirli olması için gerek ve yeter şart, (4) eşitliğini verecek köşegen değerleri pozitif bir üst üçgen R matrisinin bulunmasıdır.*

Kanıt: Önce gereklilik yönünü gösterelim. A matrisi pozitif belirli olduğu için tekil değildir. Dolayısıyla, Teorem 1'e göre A matrisinin LU ayrışımı mevcuttur. Ayrıca A matrisi simetrik de olduğuna göre (3) kullanılarak $A = LDL^\intercal$ şeklinde yazılabilir. Yine A matrisinin pozitif belirli olması nedeniyle D köşegen matrisinde $d_{ii} > 0$, $i = 1, \dots, n$ olur*. O zaman

$$D^{1/2} = \begin{bmatrix} \sqrt{d_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{d_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{d_{nn}} \end{bmatrix}$$

olacak şekilde yeni bir köşegen matris tanımlayabiliriz. Ardından

$$A = L(D^{1/2}D^{1/2})L^\intercal = \underbrace{(LD^{1/2})}_{R^\intercal}\underbrace{(LD^{1/2})^\intercal}_{R}$$

ifadesini elde ederiz. Bu eşitlikteki R matrisi, köşegen elemanları pozitif olan bir üst üçgen matristir. Son kısımda da yeterlilik yönünü kanıtlayabiliriz. Her $y \neq 0$ için

$$y^\intercal Ay = y^\intercal R^\intercal Ry = (Ry)^\intercal(Ry) = \|Ry\|^2 > 0$$

eşitsizliği elde edilir. Bu da A matrisinin pozitif belirli bir matris olduğunu gösterir. □

Dikkat ederseniz, Teorem 2'nin kanıtını Cholesky ayrışımı yapmak için bir algoritma olarak da düşünebiliriz. Yazması benden, bilgisayarda denemesi sizden.

4 Ve Ötesi

Matris ayrışımı derya gibi konu. Sırf bu iş için yazılmış kitaplar var [2]. Burada yaptığımız gibi yalnızca LU ve Cholesky ayrışımına konsantre olsak bile daha söylenecek şeyler var. Ben her iki ayrışım yöntemi için de kare matrisleri kullandım. Oysa Teorem 1'i, boyutlara dikkat ederek dikdörtgen matrisler için de pekâlâ yazabilirdik. Ayrışım konusunu anlatırken hep tersi olan matrisler kullandım. Ancak bu da şart değil. Örneğin LU ayrışımı yöntemi

* Aksi halde A matrisinin köşegen değerlerinden biri negatif olur. Alın size bir ev ödevi daha.

bir matrisin tekil olup olmadığını anlamak için de kullanılabilir.

Eğer konu hoşunuza gittiyse QR ayrışımına da bir göz atabilirsiniz [1]. Bu yazıdan sonra anlamakta zorlanmayacağınızı düşünüyorum. Ancak önceden dik matrisler konusunda bir alıştırma yapmanız yerinde olur. Asıl tekil değer ayrışımı diye bir konu var ki pek çok yerde karşınıza çıkabilir. Onun için biraz hazırlık gerek. İleride vereceğim tariflerden birini tekil değer ayrışımına ayırmayı planlıyorum.

Yazı boyunca matrislerin çarpanlarından bahsettim. Aranızda matrisler yerine tam sayıları düşünenler olacaktır. Tam sayıları asal çarpanlarına tek bir şekilde ayırabiliriz. Ancak matrislerin durumu öyle değil. Örneğin LU ayrışımı ile elde edilen çarpanlar ile QR ayrışımı ile elde edilen çarpanlar birbirlerinden farklıdır.

Bu noktada şunu da söylemeliyim. Matrisleri çarpanlarına ayırmanın tek sebebi denklem sistemlerini çözmek değil. Bir matrisin özdeğerlerini bulurken, tersini alırken ya da diğer hesaplamaları yaparken çarpanları ile çalışmak, kendisi ile çalışmaktan çok daha hızlı olabilir.

Daha önce yaptığım gibi bu yazı ve kullanılan GNU Octave işlemlerini bir sayfada topladım. Bu seferki adres şöyle:

`http://tinyurl.com/MD0517-SIB`

Bir tarifi daha bitirdik. Bu sefer doğrusal eşitliklerden bahsettik. Bir dahaki sayıda yine doğrusal denklemler ile uğraşacağız. Fakat bu sefer işin içine eşitsizlikler de girecek. Üstelik bu eşitsizliklerin tanımladığı alan üzerinde doğrusal bir fonksiyonun en küçük değerini bulacağız. Hedefimiz, doğrusal programlamanın temel taşı olan Simpleks Algoritması'nı anlamak. Hallederiz.

Kaynakça:

[1] Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM: Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[2] Gilbert W. Stewart, *Matrix Algorithms - Volume I: Basic Decompositions*, SIAM: Society for Industrial and Applied Mathematics, Philadelphia, 1998.

Veri Bilim-Yapay Öğrenme Yaz Okulu yapıldı

Üniversite öğrencileri, akademisyenler ile yapay öğrenme konusuna ilgi duyan profesyonellerin katılımıyla **Veri Bilim-Yapay Öğrenme Yaz Okulu** Boğaziçi Üniversitesi Teleiletişim ve Enformatik Teknolojileri Uygulama ve Araştırma Merkezi'nde (TETAM) gerçekleştirildi. Yaz okulunun amacı katılımcıları, veri bilim ve yapay öğrenme alanında karşılaşılan temel konuları hem kuramsal, hem de uygulamalı şekilde bilgilendirmekti.

Takip etmek için temel bir analiz, lineer cebir ve olasılık bilgisi ile başlangıç seviyesinde programlama (özellikle Python) becerisi yeterli olan yaz okulu ile ilgili bütün materyallere

`https://github.com/sibirbil/VBYO`

adresinden ulaşmak mümkün.
Öncesinde 20-23 Ekim 2016 ve 20-23 Nisan 2017 tarihleri arasında Nesin Matematik Köyü'nde düzenlenen **Yapay Öğrenmenin Matematiksel Temelleri** yaz okullarının ders notları ve uygulamaları da

`http://people.sabanciuniv.edu/sibirbil/MK/`

adresinde mevcut.
Veri bilimi ve yapay öğrenme konularında benzer okulların gelecekte de düzenlenmesi planlanıyor.