NÜMERİK ANALİZ

Bilimsel Hesaplama Matematiği, Gazi Kitabevi 2012

Nuri ÖZALP

LİNEER SİSTEMLERİN ÇÖZÜMÜ

LU ve Cholesky Ayrıştırmaları

Kolay Çözülebilir Sistemler

Amacımız

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n = b_n \end{cases}$$

formuna sahip lineer denklem sistemlerini çözmeyi nümerik bakış açısından tartışmaktır.

Matrisler denklem sistemlerini temsil etmek için kullanışlı araçlardır. Bu durumda, yukarıdaki denklem sistemini

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

şeklinde yazılabilir. Böylece bu matrisleri, denklem basitçe

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

yazılacak şekilde A, x, ve b ile gösterebiliriz.

(1) sistemi

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

köşegensel yapıda ise, bu durumda sistem n tane basit denkleme indirgenir ve çözümü de

$$x = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ b_3/a_{33} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

olur. Eğer, bazı i indisleri için $a_{ii} = 0$ ve $b_i = 0$ ise, bu durumda x_i herhangi bir reel sayı olabilir. Eğer, $a_{ii} = 0$ ve $b_i \neq 0$ ise, sistemin çözümü yoktur.

Sistem **alt üçgensel yapıda ise**, yani

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

şeklindeyse, ilk denklemden x_1 i elde ederiz. Bulunan x_1 değerini ikinci denklemde yazarak, x_2 yi çözeriz. Aynı yolla devam ederek, ardarda sırada $x_1, x_2, \ldots, x_n$ i buluruz.

Bu şekildeki formal çözüm algoritması **ileri-yönde yerleştirme** olarak adlandırılır.

girdi $n, \ (a_{ij}), (b_i)$
$i = 1$ **den** n **ye döngü**

$$x_i \leftarrow \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j \right) \Big/ a_{ii}$$

döngü sonu
çıktı (x_i)

Aynı düşünce **üst üçgensel yapıya** sahip bir sistemi çözmek için de uygulanabilir. Bu tip bir matris sistemi

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

formundadır.

Tekrar, $1 \le i \le n$ için $a_{ii} \neq 0$ kabul edilmelidir. x i çözen formal algoritma aşağıdaki gibi olup, **geri-yönde yerleştirme** olarak adlandırılır:

girdi n, $(a_{ij}), (b_i)$
$i = n$ **den** 1 **e,** -1 **adımlı döngü**
$\quad x_i \leftarrow \left(b_i - \sum_{j=i+1}^{n} a_{ij}x_j\right) / a_{ii}$
döngü sonu
çıktı (x_i)

LU-Ayrıştırmaları

A nın, bir alt üçgensel L matrisi ile bir üst üçgensel U matrisinin çarpımı şeklinde ayrıştırılabildiğini, yani $A = LU$ olduğunu kabul edelim. Bu durumda, $Ax = b$ sistemini çözmek için, iki aşamada

$$Lz = b \text{ yi } z \text{ ye göre}$$

$$Ux = z \text{ yi } x \text{ e göre}$$

çözmek yeterlidir.

- Böyle bir ayrışıma her matris sahip olmayabilir.

Bir $n \times n$ lik A matrisi ile başlayalım ve $A = LU$ olacak şekilde

$$L = \begin{bmatrix} \ell_{11} & 0 & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & \ell_{nn} \end{bmatrix} \qquad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

matrislerini araştıralım. Bunun mümkün olduğu durumlarda, A bir LU-**ayrışımına** sahiptir deriz. L ve U, Denklem (4) ten, **tek olarak** *belirlenmez*. Aslında, her i için, ℓ_{ii} veya u_{ii} den birine (fakat her ikisine değil) sıfırdan farklı bir değer *atayabiliriz*. Örneğin, basit bir seçim $i = 1, 2, ..., n$ için, $\ell_{ii} = 1$ almak ve böylece, L yi **birim alt üçgensel** yapmaktır. Bir başka açık seçenek, U yu **birim üst üçgensel** (her i için, $u_{ii} = 1$) yapmaktır. Bu iki durum özel öneme sahiptir.

A nın LU-ayrışımı için algoitma:
$s > i$ için $\ell_{is} = 0$ ve $s > j$ için $u_{sj} = 0$ olmak üzere

$$a_{ij} = \sum_{s=1}^{n} \ell_{is} u_{sj} = \sum_{s=1}^{\min(i,j)} \ell_{is} u_{sj} \tag{1}$$

İşlem sürecinin her adımı, U da yeni bir satır ve L de yeni bir kolon belirler. k-yıncı adımda U daki $1, 2, \ldots, k-1$. satırların ve L deki $1, 2, \ldots, k-1$. kolonların hesaplandığını kabul edelim. ($k = 1$ için bu kabul varsayımsal doğrudur.)
u_{kk} veya ℓ_{kk} dan biri belirlenmişse, (1) de $i = j = k$ alınırsa, diğeri

$$a_{kk} = \sum_{s=1}^{k-1} \ell_{ks} u_{sk} + \ell_{kk} u_{kk} \tag{2}$$

denkleminden elde edilir.

Böylece ℓ_{kk} ve u_{kk} bilinmek üzere, k-yıncı satır $(i = k)$ ve j. kolon $(j = k)$ için, Denklem (1) den sırası ile

$$a_{kj} = \sum_{s=1}^{k-1} \ell_{ks} u_{sj} + \ell_{kk} u_{kj} \qquad (k+1 \leq j \leq n) \tag{3}$$

$$a_{ik} = \sum_{s=1}^{k-1} \ell_{is} u_{sk} + \ell_{ik} u_{kk} \qquad (k+1 \leq i \leq n) \tag{4}$$

yazarız. Eğer $\ell_{kk} \neq 0$ ise, u_{kj} elemanlarını elde etmek için Denklem (3) kullanılabilir. Benzer olarak, eğer $u_{kk} \neq 0$ ise, ℓ_{ik} elemanlarını elde etmek için Denklem (4) kullanılabilir.

Yukarıdaki analize dayalı algoritma:

- L birim alt üçgensel ($1 \leq i \leq n$ için $\ell_{ii} = 1$) iken **Doolittle ayrıştırması**

Yukarıdaki analize dayalı algoritma:

- L birim alt üçgensel ($1 \leq i \leq n$ için $\ell_{ii} = 1$) iken **Doolittle ayrıştırması**
- U birim üst üçgensel ($1 \leq i \leq n$ için $u_{ii} = 1$) iken ise **Crout ayrıştırması** olarak bilinir.

Yukarıdaki analize dayalı algoritma:

- L birim alt üçgensel ($1 \leq i \leq n$ için $\ell_{ii} = 1$) iken **Doolittle ayrıştırması**
- U birim üst üçgensel ($1 \leq i \leq n$ için $u_{ii} = 1$) iken ise **Crout ayrıştırması** olarak bilinir.
- $U = L^T$ ve böylece $1 \leq i \leq n$ için $\ell_{ii} = u_{ii}$ iken algoritma **Cholesky ayrıştırması** olarak adlandırılır.

Yukarıdaki analize dayalı algoritma:

- L birim alt üçgensel ($1 \leq i \leq n$ için $\ell_{ii} = 1$) iken **Doolittle ayrıştırması**
- U birim üst üçgensel ($1 \leq i \leq n$ için $u_{ii} = 1$) iken ise **Crout ayrıştırması** olarak bilinir.
- $U = L^T$ ve böylece $1 \leq i \leq n$ için $\ell_{ii} = u_{ii}$ iken algoritma **Cholesky ayrıştırması** olarak adlandırılır.
- Cholesky yöntemini bu kesimin sonunda daha detaylı olarak ele alacağız, çünkü bu ayrışım A matrisinin **reel, simetrik** $(A^T = A)$ **ve pozitif tanımlı** $(x^T Ax > 0)$ olma özelliklerini gerektirmektedir.

Yukarıdaki analize dayalı algoritma:

- L birim alt üçgensel ($1 \leq i \leq n$ için $\ell_{ii} = 1$) iken **Doolittle ayrıştırması**
- U birim üst üçgensel ($1 \leq i \leq n$ için $u_{ii} = 1$) iken ise **Crout ayrıştırması** olarak bilinir.
- $U = L^T$ ve böylece $1 \leq i \leq n$ için $\ell_{ii} = u_{ii}$ iken algoritma **Cholesky ayrıştırması** olarak adlandırılır.
- Cholesky yöntemini bu kesimin sonunda daha detaylı olarak ele alacağız, çünkü bu ayrışım A matrisinin **reel, simetrik** ($A^T = A$) **ve pozitif tanımlı** ($x^T Ax > 0$) olma özelliklerini gerektirmektedir.

- Bu ayrıştırmalardan hangisi daha iyidir? Bunların herbiri basit Gauss eleme yordamının varyasyonları ile ilişkilidir. Bu nedenle, konunun tam olarak anlaşılması için bunlar hakkında bilgiye gerek vardır.

Genel LU-ayrıştırması:

girdi n, (a_{ij})

$k = 1$ **den** n **e döngü**

ℓ_{kk} veya u_{kk} dan biri için sıfırdan farklı

bir değer belirle ve diğerini hesapla

$$\ell_{kk} u_{kk} \leftarrow a_{kk} - \sum_{s=1}^{k-1} \ell_{ks} u_{sk}$$

$j = k+1$ **den** n **ye döngü**

$$u_{kj} \leftarrow \left(a_{kj} - \sum_{s=1}^{k-1} \ell_{ks} u_{sj} \right) \Big/ \ell_{kk}$$

döngü sonu

$i = k+1$ **den** n **ye döngü**

$$\ell_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is} u_{sk} \right) \Big/ u_{kk}$$

döngü sonu

döngü sonu

çıktı (ℓ_{ij}), (u_{ij})

Doolittle ayrıştırması önkodu:

girdi $n,\ (a_{ij})$

$k = 1$ **den** n **e döngü**

$\ell_{kk} \leftarrow 1$

$j = k$ **dan** n **ye döngü**

$$u_{kj} \leftarrow a_{kj} - \sum_{s=1}^{k-1} \ell_{ks} u_{sj}$$

döngü sonu

$i = k + 1$ **den** n **ye döngü**

$$\ell_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is} u_{sk} \right) \Big/ \ u_{kk}$$

döngü sonu

döngü sonu

çıktı $(\ell_{ij}), (u_{ij})$

Örnek

$$A = \begin{bmatrix} 60 & 30 & 20 \\ 30 & 20 & 15 \\ 20 & 15 & 12 \end{bmatrix}$$

matrisinin Doolittle, Crout ve Cholesky ayrıştırmalarını bulunuz.

Çözüm

Algoritmadan, Doolittle ayrışımı

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 1 & 1 \end{bmatrix} \begin{bmatrix} 60 & 30 & 20 \\ 0 & 5 & 5 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \equiv LU$$

bulunur. Diğer iki ayrışımı doğrudan hesaplamak yerine, bunları Doolittle ayrıştırmasından elde edebiliriz.

Çözüm

U nun köşegen elemanlarını bir köşegensel D matrisinde yazarak

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 1 & 1 \end{bmatrix} \begin{bmatrix} 60 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \equiv LD\hat{U}$$

buluruz. $\hat{L} = LD$ *dersek,*

$$A = \begin{bmatrix} 60 & 0 & 0 \\ 30 & 5 & 0 \\ 20 & 5 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \equiv \hat{L}\hat{U}$$

Crout ayrışımını elde ederiz.

Çözüm

Cholesky ayrışımı, $LD\hat{U}$ ayrışımında D yi $D^{1/2}D^{1/2}$ formuna ayırarak ve çarpanlardan biri L diğeri de $\hat{U}$ ile ilişkilendirerek elde edilebilir. Böylece,

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{60} & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & \frac{1}{3}\sqrt{3} \end{bmatrix} \\
&\qquad \times \begin{bmatrix} \sqrt{60} & 0 & 0 \\ 0 & \sqrt{5} & 0 \\ 0 & 0 & \frac{1}{3}\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{60} & 0 & 0 \\ \frac{1}{2}\sqrt{60} & \sqrt{5} & 0 \\ \frac{1}{3}\sqrt{60} & \sqrt{5} & \frac{1}{3}\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{60} & \frac{1}{2}\sqrt{60} & \frac{1}{3}\sqrt{60} \\ 0 & \sqrt{5} & \sqrt{5} \\ 0 & 0 & \frac{1}{3}\sqrt{3} \end{bmatrix} \equiv \hat{L}\hat{L}^T
\end{aligned}
$$

olur. ■

En çok ilgilenilen ayrışım, L bir **birim alt üçgensel** ve U da bir **üst üçgensel** olmak üzere $A = LU$ formudur.
Bir A kare matrisinin bir LU-ayrışımına sahip olması için bir yeter koşul aşağıdadır:

Teorem (LU-Ayrışımı)

Eğer $n \times n$ lik bir A kare matrisinin n tane ana temel minörleri tümüyle tekil değil ise, bu durumda A bir LU-ayrışımına sahiptir.

Teorem (LL^T-Ayrışımı için Cholesky Teoremi)

*Eğer, A reel, simetrik (yani $A = A^T$) ve pozitif tanımlı (yani $x^T Ax > 0$ bir matris ise, bu durumda L pozitif köşegenli bir alt üçgensel matris olmak üzere, A nın **tek** bir $A = LL^T$ ayrışımı vardır.*

Cholesky ayrıştırması önkodu:

girdi n, (a_{ij})
$k = 1$ **den** n **e döngü**

$$\ell_{kk} \leftarrow \left(a_{kk} - \sum_{s=1}^{k-1} \ell_{ks}^2\right)^{1/2}$$

$i = k+1$ **den** n **ye döngü**

$$\ell_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is}\ell_{ks}\right) \Big/ \ell_{kk}$$

döngü sonu
döngü sonu
çıktı (ℓ_{ij})