

第三章 线性方程

第 (2) 部分: LU 分解

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

我们的目标是解方程 $\mathbf{Ax} = \mathbf{b}$ 。高斯消元法能帮助我们同时 (1) 解方程；(2) 提供 $\mathbf{A} = \mathbf{LU}$ 分解。其中实际上资源消耗最主要的部分是在 LU 分解。

若我们把 $\mathbf{A} = \mathbf{LU}$ 作为初始点，计算 $\mathbf{LUx} = \mathbf{b}$ 可以改写成

$$\mathbf{Ly} = \mathbf{b}$$

$$\mathbf{Ux} = \mathbf{y}$$

对于任意的 n 阶、具有 LU 分解的矩阵 \mathbf{A} ，在假设 \mathbf{L} 和 \mathbf{U} 已知的情况下，仅仅计算 \mathbf{y} 和计算 \mathbf{x} 的资源消耗是：

(A) $\mathcal{O}(n)$

(B) $\mathcal{O}(n^2)$

(C) $\mathcal{O}(n^3)$

我们的目标是解方程 $\mathbf{Ax} = \mathbf{b}$ 。高斯消元法能帮助我们同时 (1) 解方程；(2) 提供 $\mathbf{A} = \mathbf{LU}$ 分解。其中实际上资源消耗最主要的部分是在 LU 分解。

若我们把 $\mathbf{A} = \mathbf{LU}$ 作为初始点，计算 $\mathbf{LUx} = \mathbf{b}$ 可以改写成

$$\mathbf{Ly} = \mathbf{b}$$

$$\mathbf{Ux} = \mathbf{y}$$

对于任意的 n 阶、具有 LU 分解的矩阵 \mathbf{A} ，在假设 \mathbf{L} 和 \mathbf{U} 已知的情况下，仅仅计算 \mathbf{y} 和计算 \mathbf{x} 的资源消耗是：

(A) $\mathcal{O}(n)$

(B) $\mathcal{O}(n^2)$

(C) $\mathcal{O}(n^3)$

答案：B

我们已经介绍了如何对于 $\mathbf{Ux} = \mathbf{y}$ 求解，类似的思想可以帮助我们写出 $\mathbf{Ly} = \mathbf{b}$ 如何求解；可参见课本的公式 (7.4.4) 和 (7.4.5)。

所以，核心的问题在于如何得到这样 LU 分解；以及对于特殊的矩阵，我们是否具有加速的算法。

$$\mathbf{A} = \mathbf{LU} = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix}.$$

目标

了解 4 种 LU 分解法

- ▶ 直接的三角分解法 (Doolittle 分解法)
- ▶ LDL 分解法
- ▶ 平方根法
- ▶ 追赶法

明白它们的适用范围和计算复杂度

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

直接的三角分解法 (Doolittle 分解法)

$$\mathbf{A} = \mathbf{LU} = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix}$$

$$= \left[\begin{array}{c|c} 1 & \\ \hline \mathbf{c} & \mathbf{L}_2 \end{array} \right] \left[\begin{array}{c|c} u_{11} & \mathbf{d}^T \\ \hline & \mathbf{U}_2 \end{array} \right] = \left[\begin{array}{c|c} u_{11} & \mathbf{d}^T \\ \hline u_{11}\mathbf{c} & \mathbf{cd}^T + \mathbf{L}_2\mathbf{U}_2 \end{array} \right]$$

通过对比元素，我们可知 $u_{11} = a_{11}$, $\mathbf{d}^T = (a_{12} \ a_{13} \ \cdots \ a_{1n})$, 并且 $\mathbf{c}^T = \frac{1}{u_{11}} [a_{21} \ a_{31} \ \cdots \ a_{n1}]$ 。

由于我们已经知道 \mathbf{c} 和 \mathbf{d} , 因而

$$\mathbf{L}_2 \mathbf{U}_2 = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} - \mathbf{c} \mathbf{d}^T$$

该计算中分别需要_____次乘法和减法。

通过对比元素，我们可知 $u_{11} = a_{11}$, $\mathbf{d}^T = (a_{12} \ a_{13} \ \cdots \ a_{1n})$, 并且 $\mathbf{c}^T = \frac{1}{u_{11}} [a_{21} \ a_{31} \ \cdots \ a_{n1}]$ 。

由于我们已经知道 \mathbf{c} 和 \mathbf{d} , 因而

$$\mathbf{L}_2 \mathbf{U}_2 = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} - \mathbf{c} \mathbf{d}^T$$

该计算中分别需要 $(n-1)^2$ 次乘法和减法。

对于进一步的得到 \mathbf{L}_2 和 \mathbf{U}_2 的过程，我们继续重复上述的步骤。

例子

请计算如下矩阵的 LU 分解:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

例子

请计算如下矩阵的 LU 分解:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

答案: $\begin{bmatrix} 1 & & \\ 2 & 1 & \\ 3 & -5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ & 1 & -4 \\ & & -24 \end{bmatrix}$

在用计算机计算时,由于计算好 u_{ri} 后 a_{ri} 就不用了,因此计算好 L,U 的元素后就存放在 A 的相应位置.例如

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \rightarrow \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ l_{21} & u_{22} & u_{23} & u_{24} \\ l_{31} & l_{32} & u_{33} & u_{34} \\ l_{41} & l_{42} & l_{43} & u_{44} \end{pmatrix}.$$

最后在存放 A 的数组中得到 L,U 的元素.

该方法使用的乘除法次数是 $\sum_{i=2}^n i(i-1) \approx \frac{n^3}{3} = \mathcal{O}(n^3)$ 。

和高斯消元法是一样的; 此处介绍了一个巧妙地能节约计算机内存的方法。

上述的方法可以等价改写成如下的算法：

Algorithm 1: LU decomposition

- 1 $u_{1i} = a_{1i}, (i = 1, 2, \dots, n)$
 - 2 $l_{i1} = a_{i1}/u_{11} (i = 2, 3, \dots, n)$
 - 3 **for** $r \leftarrow 2$ **to** n **do**
 - 4 $u_{ri} = a_{ri} - \sum_{k=1}^{r-1} l_{rk}u_{ki} (i = r, r+1, \dots, n)$
 - 5 $l_{ir} = (a_{ir} - \sum_{k=1}^{r-1} l_{ik}u_{kr})/u_{rr} (i = r+1, \dots, n, r \neq n)$
 - 6 **end**
-

(大家可以根据自己偏好选择不同视角理解 Doolittle 分解法)

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

LDL 法

假设 \mathbf{A} 具有对称性，我们能否利用对称性来设计加速算法。

定理 (对称矩阵的三角分解定理)

假设 \mathbf{A} 是 n 阶矩阵，且 \mathbf{A} 的所有顺序主子式均不等于 0（其实目的就是保证存在 LU 分解），则 \mathbf{A} 可以唯一分解为

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T$$

其中 \mathbf{L} 是单位下三角矩阵， \mathbf{D} 是对角矩阵。

(证明请见课内板书或课本)

问题变成了如何找到 L 和 D 使得

$$\mathbf{A} = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & & \\ & d_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & d_n \end{pmatrix} \begin{pmatrix} 1 & l_{21} & l_{31} & \cdots & l_{n1} \\ & 1 & l_{32} & \cdots & l_{n2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & l_{n,n-1} \\ & & & & 1 \end{pmatrix}$$

问题：我们可以怎么做？

$$\mathbf{A} = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & & \\ & d_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & d_n \end{pmatrix} \begin{pmatrix} 1 & l_{21} & l_{31} & \cdots & l_{n1} \\ & 1 & l_{32} & \cdots & l_{n2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & l_{n,n-1} \\ & & & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ \vdots & \ddots & \ddots & & \\ l_{n1} & \cdots & l_{n,n-1} & & 1 \end{pmatrix} \begin{pmatrix} d_1 & d_1 l_{21} & \cdots & d_1 l_{n1} \\ & d_2 & \cdots & d_2 l_{n2} \\ & & \ddots & \vdots \\ & & & d_n \end{pmatrix}$$

$$= \left[\begin{array}{c|ccc} d_1 & d_1 l_{21} & \cdots & d_1 l_{n1} \\ \hline d_1 l_{21} & \mathbf{L}_2 \mathbf{U}_2 + d_1 \begin{bmatrix} l_{21} \\ \vdots \\ l_{n1} \end{bmatrix} & [l_{21} & \cdots & l_{n1}] \\ \vdots & & & \\ d_1 l_{n1} & & & \end{array} \right]$$

通过对照，我们需要

$$\begin{aligned}d_1 &= a_{11} \\ l_{i1} &= \frac{a_{i1}}{d_1} \quad i = 2, \dots, n\end{aligned}$$

下一步计算

$$\mathbf{L}_2 \mathbf{U}_2 = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} - d_1 \begin{bmatrix} l_{21} \\ \vdots \\ l_{n1} \end{bmatrix} [l_{21} \quad \cdots \quad l_{n1}]$$

我们通过类似直接三角分解法，可以改写出课本公式 (7.4.9)。

我们来考虑一下乘除法运算的次数（对于至此为止的第一层的计算）。

对于至此为止的第一层的计算，乘除法运算次数为

$$\underbrace{(n-1)}_{\text{第一列}} + \underbrace{\frac{(n-1)^2 - (n-1)}{2} + (n-1)}_{L_2 U_2 \text{ 这一分块矩阵}} \approx \frac{n^2}{2}$$

因而，总的计算量大约为 $\frac{n^3}{6}$ ，是高斯消元法的一半。

问题：此处消耗只有一半的来源是什么？

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

平方根法

平方根法主要是针对（半）正定矩阵。

定义： 假设 \mathbf{A} 是对称矩阵，

- ▶ 对于任意的向量 \mathbf{v} ， $\mathbf{v}^T \mathbf{A} \mathbf{v} \geq 0$ ，则我们称该矩阵是半正定矩阵。
- ▶ 对于任意的非零向量 \mathbf{v} ， $\mathbf{v}^T \mathbf{A} \mathbf{v} > 0$ ，则我们称该矩阵是正定矩阵。

结论：

- (1) 若我们取 $\mathbf{v} = \mathbf{e}_i$ ，则我们可以得到 $\mathbf{v}^T \mathbf{A} \mathbf{v} = a_{ii} \geq 0$ 。
- (2) 若 \mathbf{A} 是半正定矩阵，且 $\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{L}^T$ ，则我们可知对角矩阵 \mathbf{D} 的元素皆为非负。

$$\mathbf{D} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} = \begin{pmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_n} \end{pmatrix} \begin{pmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_n} \end{pmatrix} = \mathbf{D}^{\frac{1}{2}} \mathbf{D}^{\frac{1}{2}}$$

因而,

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T = \mathbf{L}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}\mathbf{L}^T = \left(\mathbf{L}\mathbf{D}^{\frac{1}{2}}\right) \left(\mathbf{L}\mathbf{D}^{\frac{1}{2}}\right)^T = \mathbf{L}_1\mathbf{L}_1^T$$

其中 $\mathbf{L}_1 = \mathbf{L}\mathbf{D}^{\frac{1}{2}}$ 是下三角矩阵。

定理 (Cholesky 分解)

如果 \mathbf{A} 为 n 阶对称正定矩阵, 则存在一个实的非奇异下三角阵 \mathbf{L} 使 $\mathbf{A} = \mathbf{L}\mathbf{L}^T$, 当限定 \mathbf{L} 的对角元素为正时, 这种分解是唯一的。

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ & l_{22} & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & l_{nn} \end{pmatrix}$$

$$= \left[\begin{array}{c|c} l_{11} & \\ \hline \mathbf{c} & \mathbf{L}_2 \end{array} \right] \left[\begin{array}{c|c} l_{11} & \mathbf{c}^T \\ \hline & \mathbf{L}_2^T \end{array} \right] = \left[\begin{array}{c|c} l_{11}^2 & l_{11}\mathbf{c}^T \\ \hline l_{11}\mathbf{c} & \mathbf{c}\mathbf{c}^T + \mathbf{L}_2\mathbf{L}_2^T \end{array} \right]$$

因此, $l_{11} = \sqrt{a_{11}}$, $l_{j1} = \frac{a_{j1}}{\sqrt{a_{11}}}$ $j = 2, 3, \dots, n$ 。剩余部分的更新实际上和前面没有区别; 因此此处略去。

该分解需要计算根号, 除此之外, 其所需的乘除法运算次数和 LDL 方法基本相同。

定理

对于任意的 $k \leq j$ ，我们可知

$$l_{jk}^2 \leq a_{jj} \leq \max_{1 \leq j \leq n} a_{jj}$$

因而该分解产生的下三角矩阵的任何元素的平方都不超过 $\max_{1 \leq j \leq n} \{a_{jj}\}$ 。

证明.

我们仅需验证通过对比 $a_{jj} = \sum_{k=1}^j l_{jk}^2$ 。



含义：该结论是说明对于 Cholesky 分解，分解得到的下三角矩阵的元素的数量级不会增加。

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

$$\begin{aligned}
\mathbf{A} &= \begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix} \\
&= \begin{pmatrix} \alpha_1 & & & & \\ \gamma_2 & \alpha_2 & & & \\ & \gamma_3 & \alpha_3 & & \\ & & \ddots & \ddots & \\ & & & \gamma_n & \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \beta_1 & & & \\ & 1 & \beta_2 & & \\ & & \ddots & \ddots & \\ & & & 1 & \beta_{n-1} \\ & & & & 1 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_1 & \alpha_1\beta_1 & & & \\ \gamma_2 & \gamma_2\beta_1 + \alpha_2 & \alpha_2\beta_2 & & \\ & \gamma_3 & \gamma_3\beta_2 + \alpha_3 & \alpha_3\beta_3 & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots \end{pmatrix}
\end{aligned}$$

所有的 $\gamma_2, \gamma_2, \dots, \gamma_n$ 可以直接得到。然后计算顺序是 $\alpha_1, \beta_1, \alpha_2, \beta_2, \dots$ 。

总结一下求解 LU 分解中待定系数的顺序

- ▶ 直接写答案: $\gamma_i = a_i$,
 $i = 2, \dots, n$

- ▶ 第一行的结果: $\alpha_1 = b_1$,

$$\beta_1 = \frac{c_1}{\alpha_1}$$

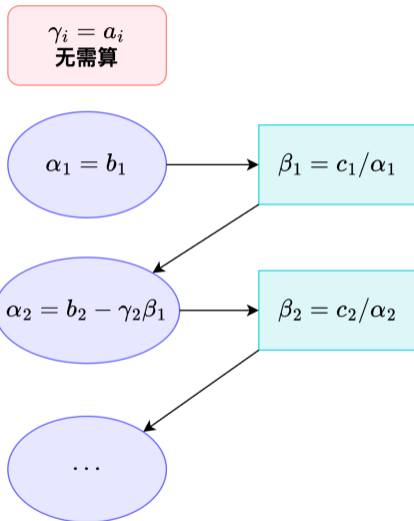
- ▶ 后面行的结果: for
 $i = 2, 3, \dots, n-1$, 令

$$\alpha_i = b_i - \gamma_i \beta_{i-1}$$

$$\beta_i = c_i / \alpha_i$$

- ▶ 最后一行的结果:

$$\alpha_n = b_n - \gamma_n \beta_{n-1}$$



最终解方程 $\mathbf{Ax} = \mathbf{f}$ 的过程变成了

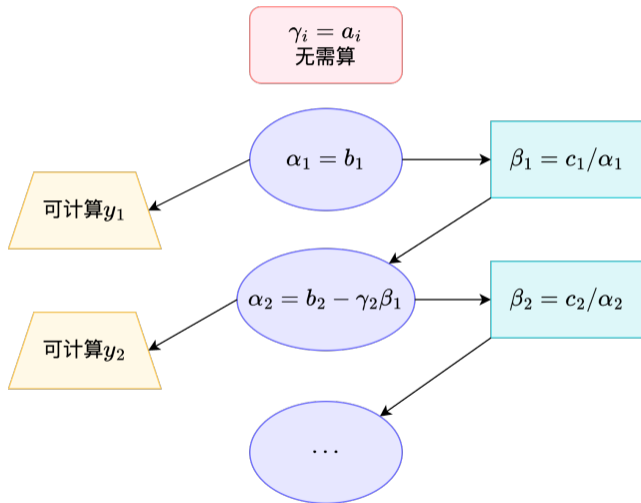
$$\mathbf{Ly} = \mathbf{f}$$

$$\mathbf{Ux} = \mathbf{y}$$

我们先来看 $\mathbf{Ly} = \mathbf{f}$,

$$\begin{pmatrix} \alpha_1 & & & & & \\ \gamma_2 & \alpha_2 & & & & \\ & \gamma_3 & \alpha_3 & & & \\ & & \ddots & \ddots & & \\ & & & \gamma_n & \alpha_n & \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_n \end{pmatrix}$$

追的过程



赶的过程

对于解 $Ux = y$,

$$\begin{pmatrix} 1 & \beta_1 & & & & \\ & 1 & \beta_2 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & \beta_{n-1} & \\ & & & & 1 & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{pmatrix}$$

$$x_n = y_n$$

$$x_{n-1} = -\beta_{n-1}x_n + y_{n-1}$$

$$x_{n-2} = -\beta_{n-2}x_{n-1} + y_{n-2}$$

\vdots

练习：对于追赶法的计算量是多少？

(A) $\mathcal{O}(n)$

(B) $\mathcal{O}(n^2)$

(C) $\mathcal{O}(n^3)$

(答案可见最后一页的总结。)

定理 (定理 7.9)

设有三对角方程组 $\mathbf{Ax} = \mathbf{f}$ ，其中 \mathbf{A} 满足对角占优条件，则 \mathbf{A} 为非奇异矩阵且由追赶法计算公式中 $\{\alpha_i\}, \{\beta_i\}$ 满足：

$$1^\circ \quad 0 < |\beta_i| < 1 \quad (i = 1, 2, \dots, n-1)$$

$$2^\circ \quad 0 < |c_i| \leq |b_i| - |a_i| < |\alpha_i| < |b_i| + |a_i| \quad (i = 2, 3, \dots, n-1)$$

$$0 < |b_n| - |a_n| < |\alpha_n| < |b_n| + |a_n|$$

目录

1. 概览
2. 直接的三角分解法
3. LDL 法
4. 平方根法
5. 追赶法
6. 总结

我们假设矩阵 $\mathbf{A} = \mathbf{LU}$ 存在 LU 分解。

| 方法 | 公式 | 乘除法次数 | 额外假设 |
|--------------------|-------------------------------|-------------------|---|
| 直接三角分解法 | $\mathbf{A} = \mathbf{LU}$ | 约 $\frac{n^3}{3}$ | 无 |
| LDL | $\mathbf{A} = \mathbf{LDL}^T$ | 约 $\frac{n^3}{6}$ | \mathbf{A} 是对称矩阵 |
| 平方根法 (Cholesky 分解) | $\mathbf{A} = \mathbf{LL}^T$ | 约 $\frac{n^3}{6}$ | \mathbf{A} 是正定 |
| 追赶法 | $\mathbf{A} = \mathbf{LU}$ | $\mathcal{O}(n)$ | \mathbf{A} 是对角占优 \mathbf{A} 为三对角矩阵 |