Matrix Computations
Lecture 2: Least-squares Problems & QR Decomposition

Tianqi Xu
Supervisors: Jie Lu
ShanghaiTech University

Main content

- Least-squares Problems and Applications
- QR Decomposition

Least-squares Problem and Its Solution

LS Problem: Given $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{y} \in \mathbb{R}^m$, solve

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \qquad \text{(LS)}$$

Theorem (LS Optimality Condition)

$\mathbf{x}_{\mathrm{LS}} \in \mathbb{R}^n$ is an optimal solution to the LS problem $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ if and only if it satisfies the following normal equation:

$$\mathbf{A}^T \mathbf{A} \mathbf{x}_{\mathrm{LS}} = \mathbf{A}^T \mathbf{y}. \qquad (*)$$

- When $\mathbf{A}$ has full-column rank, the *unique* solution to $(*)$ is

$$\mathbf{x}_{\mathrm{LS}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$$

Application 1: Polynomial approximation

Aim: Given a set of input-output data pairs $(t_i, y_i) \in \mathbb{R} \times \mathbb{R}$, $i = 1, \ldots, N$, find a polynomial $p(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_p t^p$ that approximates the data by minimizing the residual $\sum_{i=1}^{N} (y_i - p(t_i))^2$, where $p \leq N$.

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_t \\ \vdots \\ y_N \end{bmatrix}}_{=\mathbf{y}} = \underbrace{\begin{bmatrix} 1 & t_1 & \cdots & t_1^p \\ 1 & t_2 & \cdots & t_2^p \\ \vdots & & & \vdots \\ 1 & t_N & \cdots & t_N^p \end{bmatrix}}_{=\mathbf{A}(\text{full rank for distinct } t_i)} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix}}_{=\mathbf{x}}$$

The polynomial approximation problem can be viewed as

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \longrightarrow \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$$

Exercise: Linear prediction

Aim: Predict future values of a time-series based on linear prediction. Suppose y_t is determined by $y_{t-1}, \ldots, y_{t-m}$ at time t, i.e.,

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_m y_{t-m}, \quad t = 1, 2, \ldots$$

Now given $y_0, y_1, \ldots, y_{99}$ $(n \geq m)$, predict $y_{100}, y_{101}, \ldots, y_{199}$ with predictor coefficients a_1, a_2, a_3, a_4.

hint: model the linear prediction as a least-square problem

Application 2: System identification

System identification: Given an input signal block $\{x_t\}_{t=0}^{T-1}$ and an output signal block $\{y_t\}_{t=0}^{T-1}$, find the system impulse response $\{h_t\}_{t=0}^{p}$

- Applications: channel estimation in communications, identification of acoustic impulse responses, etc.

$$\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_p \\ \vdots \\ \vdots \\ y_{T-1} \end{bmatrix}}_{=\mathbf{y}} = \underbrace{\begin{bmatrix} x_0 & & & \\ x_1 & x_0 & & \\ \vdots & & \ddots & \\ x_p & \cdots & x_1 & x_0 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ x_{T-1} & \cdots & x_{T-p} & x_{T-1-p} \end{bmatrix}}_{=\mathbf{X}} \underbrace{\begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_p \end{bmatrix}}_{=\mathbf{h}}$$

The system impulse response $\mathbf{h}$ can estimated by

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{y} - \mathbf{X}\mathbf{h}\|_2^2 \longrightarrow \mathbf{h} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

Application 3: Smoothing

Aim: Find a smooth signal $\mathbf{x}$ that approximates a noisy signal $\mathbf{y}$.
Method: One approach to smooth a noisy signal is based on least squares weighted regularization.

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda\|\mathbf{D}\mathbf{x}\|^2 \tag{1}$$

- $\lambda > 0$ a parameter for tuning smoothness. Lager λ, more smooth signal $\mathbf{x}$.
- $\mathbf{D} = \begin{bmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & & \ddots & & \\ & & & & 1 & -2 & 1 \end{bmatrix}$, $\mathbf{D}\mathbf{x}$ is a discrete form of the second-order derivative of $\mathbf{x}$, which measures the smoothness of a signal. the signal x(n).
- The solution of (1) is $\mathbf{x} = \left(\mathbf{I} + \lambda\mathbf{D}^\top\mathbf{D}\right)^{-1}\mathbf{y}$ (setting the derivative to zero).

Exercise: Deconvolution

Deconvolution: Given an output signal block $\{y_t\}_{t=0}^{T-1}$ and the system impulse response $\{h_t\}_{t=0}^{p}$, estimate the input signal block $\{x_t\}_{t=0}^{T-1}$

$$\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_p \\ \vdots \\ \vdots \\ y_{T-1} \end{bmatrix}}_{=\mathbf{y}} = \underbrace{\begin{bmatrix} h_0 & & & & & & \\ h_1 & h_0 & & & & & \\ \vdots & & \ddots & & & & \\ h_p & \dots & h_1 & h_0 & & & \\ & \ddots & & & \ddots & & \\ & & \ddots & & & \ddots & \\ & & & h_p & \dots & h_1 & h_0 \end{bmatrix}}_{=\mathbf{A}\in\mathbb{R}^{T\times T}} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_p \\ \vdots \\ \vdots \\ x_{T-1} \end{bmatrix}}_{=\mathbf{x}}$$

- Since $\mathbf{A}$ is often singular, instead of solving $\mathbf{y} = \mathbf{A}\mathbf{x}$, we solve the problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|^2 \tag{2}$$

Aim: Using the given data, plot the input signal $\mathbf{x}$ under different λ.

QR decomposition

Thin QR Decomposition for Full Column-Rank Matrices

Suppose $\mathbf{A} \in \mathbb{R}^{m \times n}$ has full column rank. Then, $\mathbf{A}$ admits a decomposition

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R_1} \qquad \text{(Thin QR Decomposition)}$$

- $\mathbf{Q}_1 \in \mathbb{R}^{m \times n}$ is semi-orthogonal
- $\mathbf{R_1} \in \mathbb{R}^{\mathbf{n} \times \mathbf{n}}$ is nonsingular and upper triangular

Gram-Schmidt Procedure (cont'd)

Algorithm: Gram-Schmidt
input: a collection of linearly independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$
$\tilde{\mathbf{q}}_1 = \mathbf{a}_1$, $\mathbf{q}_1 = \tilde{\mathbf{q}}_1 / \|\tilde{\mathbf{q}}_1\|_2$
for $i = 2, \ldots, n$
  $\tilde{\mathbf{q}}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} (\mathbf{q}_j^T \mathbf{a}_i)\mathbf{q}_j$
  $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|_2$
end
output: $\mathbf{q}_1, \ldots, \mathbf{q}_n$

- Complexity: $\mathcal{O}(mn^2)$
- $\mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\} = \mathrm{span}\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n\}$
- $\begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$ is a semi-orthogonal matrix

Thin QR Decomposition via Gram-Schmidt

From Gram-Schmidt,

$$\mathbf{a}_i = \sum_{j=1}^{i} r_{ji}\mathbf{q}_j, \quad i = 1, \ldots, n$$

where

$$r_{ii} = \|\tilde{\mathbf{q}}_i\|_2, \quad r_{ji} = \mathbf{q}_j^T \mathbf{a}_i, \; j = 1, \ldots, i-1$$

Equivalently,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \mathbf{Q}_1\mathbf{R}_1$$

where

- $\mathbf{Q}_1 = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$ semi-orthogonal
- $\mathbf{R}_1$ is nonsingular, upper triangular with $[\mathbf{R}_1]_{\mathbf{ij}} = \mathbf{r}_{\mathbf{ij}}$ for $i \leq j$

Example & Exercise

- **Example 1**: Use the Gram-Schmidt to decompose

$$\mathbf{A} = \begin{bmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{bmatrix}$$

- **Exercise 1**: Use the Gram-Schmidt to decompose

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

- *see Chapter 5.2.7 of* **[Golub-van-Loan'13]** *for the Gram-Schmidt Algorithm*

Numerical Error Issue of Gram-Schmidt

Gram-Schmidt is numerically unstable due to propagation of numerical errors

Example: Given
$\mathbf{a}_1 = \begin{bmatrix} 1 & \epsilon & 0 & 0 \end{bmatrix}^T, \mathbf{a}_2 = \begin{bmatrix} 1 & 0 & \epsilon & 0 \end{bmatrix}^T, \mathbf{a}_3 = \begin{bmatrix} 1 & 0 & 0 & \epsilon \end{bmatrix}^T$ with tiny ϵ so that the approximation $1 + \epsilon^2 \approx 1$ can be made

Applying Gram-Schmidt with the above approximation yields

- $\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|_2} = \begin{bmatrix} 1 & \epsilon & 0 & 0 \end{bmatrix}^T$
- $\tilde{\mathbf{q}}_2 = \mathbf{a}_2 - \mathbf{q}_1^T \mathbf{a}_2 \mathbf{q}_1 = \begin{bmatrix} 0 & -\epsilon & \epsilon & 0 \end{bmatrix}^T$
  $\mathbf{q}_2 = \frac{\tilde{\mathbf{q}}_2}{\|\tilde{\mathbf{q}}_2\|_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 1 & 0 \end{bmatrix}^T$
- $\tilde{\mathbf{q}}_3 = \mathbf{a}_3 - \mathbf{q}_1^T \mathbf{a}_3 \mathbf{q}_1 - \mathbf{q}_2^T \mathbf{a}_3 \mathbf{q}_2 = \begin{bmatrix} 0 & -\epsilon & 0 & \epsilon \end{bmatrix}^T$
  $\mathbf{q}_3 = \frac{\tilde{\mathbf{q}}_3}{\|\tilde{\mathbf{q}}_3\|_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 & 1 \end{bmatrix}^T$

  Orthogonality is lost!

Modified Gram-Schmidt

Instead of computing $\tilde{\mathbf{q}}_i = \mathbf{a}_i - (\mathbf{q}_1^T \mathbf{a}_i)\mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_i)\mathbf{q}_2 - \cdots - (\mathbf{q}_{i-1}^T \mathbf{a}_i)\mathbf{q}_{i-1}$ in Gram-Schmidt (full column rank case), compute

$$
\begin{aligned}
\tilde{\mathbf{q}}_i^{(1)} =& \mathbf{a}_i - (\mathbf{q}_1^T \mathbf{a}_i)\mathbf{q}_1 \\
\tilde{\mathbf{q}}_i^{(2)} =& \tilde{\mathbf{q}}_i^{(1)} - (\mathbf{q}_2^T \tilde{\mathbf{q}}_i^{(1)})\mathbf{q}_2 \\
&\vdots \\
\tilde{\mathbf{q}}_i^{(j)} =& \tilde{\mathbf{q}}_i^{(j-1)} - (\mathbf{q}_j^T \tilde{\mathbf{q}}_i^{(j-1)})\mathbf{q}_j \\
&\vdots \\
\tilde{\mathbf{q}}_i = \tilde{\mathbf{q}}_i^{(i-1)} =& \tilde{\mathbf{q}}_i^{(i-2)} - (\mathbf{q}_{i-1}^T \tilde{\mathbf{q}}_i^{(i-2)})\mathbf{q}_{i-1}
\end{aligned}
$$

Modified Gram-Schmidt (cont'd)

Example (revisit): Given
$\mathbf{a}_1 = \begin{bmatrix} 1 & \epsilon & 0 & 0 \end{bmatrix}^T, \mathbf{a}_2 = \begin{bmatrix} 1 & 0 & \epsilon & 0 \end{bmatrix}^T, \mathbf{a}_3 = \begin{bmatrix} 1 & 0 & 0 & \epsilon \end{bmatrix}^T$ with tiny ϵ
so that the approximation $1 + \epsilon^2 \approx 1$ can be made

Applying modified Gram-Schmidt with the above approximation yields

- $\tilde{\mathbf{q}}_1 = \begin{bmatrix} 1 & \epsilon & 0 & 0 \end{bmatrix}^T$
  $\mathbf{q}_1 = \begin{bmatrix} 1 & \epsilon & 0 & 0 \end{bmatrix}^T$
- $\tilde{\mathbf{q}}_2 = \mathbf{a}_2 - \mathbf{q}_1^T \mathbf{a}_2 \mathbf{q}_1 = \begin{bmatrix} 0 & -\epsilon & \epsilon & 0 \end{bmatrix}^T$
  $\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 1 & 0 \end{bmatrix}^T$
- $\tilde{\mathbf{q}}_3^{(1)} = \mathbf{a}_3 - \mathbf{q}_1^T \mathbf{a}_3 \mathbf{q}_1 = \begin{bmatrix} 0 & -\epsilon & 0 & \epsilon \end{bmatrix}^T$
  $\tilde{\mathbf{q}}_3 = \tilde{\mathbf{q}}_3^{(2)} = \tilde{\mathbf{q}}_3^{(1)} - \mathbf{q}_2^T \tilde{\mathbf{q}}_3^{(1)} \mathbf{q}_2 = \begin{bmatrix} 0 & -\frac{\epsilon}{2} & -\frac{\epsilon}{2} & \epsilon \end{bmatrix}^T$
  $\mathbf{q}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & -1 & -1 & 2 \end{bmatrix}^T$

  Orthogonality is preserved approximately

Example & Exercise

- **Example 2**: Use the modified Gram-Schmidt to decompose

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0.001 & 0.001 & 0 \\ 0.001 & 0 & 0.001 \end{bmatrix}$$

- **Exercise 2**: Use the modified Gram-Schmidt to decompose

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0.001 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{bmatrix}$$

  - *see Chapter 5.2.8 of* **[Golub-van-Loan'13]** *for the modified Gram-Schmidt Algorithm*

Householder Reflections

Householder reflection: Given $\mathbf{x} \in \mathbb{R}^m$, let

$$\mathbf{v} = \mathbf{x} \mp \|\mathbf{x}\|_2 \mathbf{e}_1 \qquad \mathbf{H} = \mathbf{I} - \frac{2}{\|\mathbf{v}\|_2^2} \mathbf{v}\mathbf{v}^T$$

$$\Longrightarrow \quad \mathbf{H}\mathbf{x} = \pm \begin{bmatrix} \|\mathbf{x}\|_2 \\ \mathbf{0} \end{bmatrix} = \|\mathbf{x}\|_2 \mathbf{e}_1$$

- $\mathbf{H} \in \mathbb{R}^{\mathbf{m} \times \mathbf{m}}$ is orthogonal
- The sign in the expression of $\mathbf{v}$ may be determined to be the one that maximizes $\|\mathbf{v}\|_2$ for the sake of numerical stability

Householder QR

1. Let $\mathbf{H_1} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}}$ be the Householder reflection w.r.t. $\mathbf{a}_1$. Transform $\mathbf{A}$ as

$$\mathbf{A}^{(1)} = \mathbf{H_1}\mathbf{A} = \begin{bmatrix} \times & \times & \dots & \times \\ 0 & {\color{red}\times} & \dots & \times \\ \vdots & \vdots & & \vdots \\ 0 & {\color{red}\times} & \dots & \times \end{bmatrix}$$

2. Let $\tilde{\mathbf{H}}_2 \in \mathbb{R}^{(m-1)\times(m-1)}$ be the Householder reflection w.r.t. $\mathbf{A}^{(1)}(2:m,2)$ (marked red above). Transform $\mathbf{A}^{(1)}$ as

$$\mathbf{A}^{(2)} = \underbrace{\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_2 \end{bmatrix}}_{=\mathbf{H}_2} \mathbf{A}^{(1)} = \begin{bmatrix} \times & \times \quad \dots \times \\ \mathbf{0} & \tilde{\mathbf{H}}_2\mathbf{A}^{(1)}(2:m,2:n) \end{bmatrix} = \begin{bmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ \vdots & 0 & \times & \dots & \times \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \times & \dots & \times \end{bmatrix}$$

- $\mathbf{A}^{(n-1)} = \mathbf{H_{n-1}} \cdots \mathbf{H_2 H_1 A}$, $\quad \mathbf{A}^{(n-1)}$ is upper triangular
- $\mathbf{R} = \mathbf{A}^{(\mathbf{n-1})}$ and $\mathbf{Q} = (\mathbf{H_{n-1}} \cdots \mathbf{H_2 H_1})^{\mathbf{T}}$

Example & Exercise

- **Example 3**: Use the Householder reflection to decompose

$$\mathbf{A} = \begin{bmatrix} 1 & 19 & -34 \\ -2 & -5 & 20 \\ 2 & 8 & 37 \end{bmatrix}$$

- **Exercise 3**: Use the Householder reflection to decompose

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

- *see Algorithm 5.2.1 in* **[Golub-van-Loan'13]** *for the Householder QR Algorithm*

Givens Rotations

Example: Let

$$\mathbf{J} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

where $c = \cos(\theta), s = \sin(\theta)$ for some θ

$$\mathbf{y} = \mathbf{J}\mathbf{x} \Longleftrightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} cx_1 + sx_2 \\ -sx_1 + cx_2 \end{bmatrix}$$

Observe that

- $\mathbf{J}$ is orthogonal
- $y_2 = 0$ if $\theta = \tan^{-1}(x_2/x_1)$, or if

$$c = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad s = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}$$

Givens QR

Example: Consider a 4×3 matrix.

$$\mathbf{A} = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,2}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,3}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

$$\xrightarrow{\mathbf{J}_{2,3}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{2,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{3,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{R}$$

- $\mathbf{B} \xrightarrow{\mathbf{J}} \mathbf{C}$ means $\mathbf{C} = \mathbf{JB}$
- **Givens QR** $(m \geq n)$: Perform a sequence of Givens rotations to annihilate the lower triangular parts of $\mathbf{A}$

$$\underbrace{(\mathbf{J}_{n,m} \dots \mathbf{J}_{n,n+2}\mathbf{J}_{n,n+1}) \dots (\mathbf{J}_{2m} \dots \mathbf{J}_{24}\mathbf{J}_{23})(\mathbf{J}_{1m} \dots \mathbf{J}_{13}\mathbf{J}_{12})}_{=\mathbf{Q}^T} \mathbf{A} = \mathbf{R}$$

where $\mathbf{R}$ is upper triangular and $\mathbf{Q}$ is orthogonal

Example & Exercise

- **Example 4**: Use the Givens rotations to decompose

$$\mathbf{A} = \begin{bmatrix} 1 & 19 & -34 \\ -2 & -5 & 20 \\ 2 & 8 & 37 \end{bmatrix}$$

- **Exercise 3**: Use the Givens rotations to decompose

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

- *see Algorithm 5.2.4 in* **[Golub-van-Loan'13]** *for the Givens QR Algorithm*

References

[Golub-van-Loan'13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4rd edition, JHU Press, 2013.