CS 3331 Numerical Methods

Lecture 4: QR decomposition

Cherung Lee

Outline

- Linear algebra
- QR factorization
  - Gram-Schmidt process
  - Householder decomposition
  - Givens rotation

Linear Algebra

Linear transformation

- A vector $\mathbf{v} = [v_1, v_2, \cdots, v_n]^T$ can be viewed as

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots v_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

- A matrix $\mathbf{A}$ can be viewed as a linear transformation.
  - Let $[\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$ be the column vectors of $\mathbf{A}$.
  - $\mathbf{A}\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + \cdots + v_n\mathbf{a}_n$ is a linear transformation of vector $\mathbf{v}$.

Orthogonal matrix LVF pp.11,160

- Two vectors $\mathbf{a}, \mathbf{b}$ are said *orthogonal* if their inner product equals to zero, $\mathbf{a}^T\mathbf{b} = 0$.
  - They are said *orthonormal* if $\mathbf{a}^T\mathbf{b} = 0$ and $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$.
- A matrix $\mathbf{Q}$ is *orthogonal* if $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$.
  - Let $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n]$, where $\mathbf{q}_i^T\mathbf{q}_j = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}$
  - $\mathbf{Q}^{-1} = \mathbf{Q}^T$ is also an orthogonal matrix.
  - $\|\mathbf{Q}\| = 1$ and $\|\mathbf{Q}\mathbf{v}\| = \|\mathbf{v}\|$
    Proof: For 2-norm, $\|\mathbf{Q}\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{Q}^T\mathbf{Q}\mathbf{v} = \mathbf{v}^T\mathbf{I}\mathbf{v} = \|\mathbf{v}\|^2$.

QR decomposition LVF pp.160

- A matrix $\mathbf{A}$ can be expressed as the product of an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$,

$$(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n) = (\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n) \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix}$$

$$\begin{aligned} \mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\ \mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\ \vdots &= \vdots \\ \mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n \end{aligned}$$

QR Factorization

How to compute QR factorization

- Gram-Schmidt process (using projection matrix)

- Householder method (using reflection matrix)

- Givens rotation (using rotation matrix)

Method I. Gram-Schmidt process

- If $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2)$, then
  - $\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$, $r_{11} = \|\mathbf{a}_1\|$.
  - Want $\mathbf{q}_2$, r_{12} and r_{22} $\begin{cases} \mathbf{q}_1^T \mathbf{q}_2 = 0; \\ \mathbf{q}_2^T \mathbf{q}_2 = 1; \\ \mathbf{a}_2 = r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2. \end{cases}$
  - Let $\mathbf{q}_2 = r_{22}^{-1}(\mathbf{a}_2 - r_{12}\mathbf{q}_1)$.

$$\begin{aligned} \mathbf{q}_1^T \mathbf{q}_2 &= r_{22}^{-1}(\mathbf{q}_1^T \mathbf{a}_2 - r_{12}\mathbf{q}_1^T \mathbf{q}_1) = r_{22}^{-1}(\mathbf{q}_1^T \mathbf{a}_2 - r_{12}) = 0 \\ r_{12} &= \mathbf{q}_1^T \mathbf{a}_2. \\ r_{22} &= \|\mathbf{a}_2 - r_{12}\mathbf{q}_1\| \text{ (so that } \|q_2\| = 1) \end{aligned}$$

The similar idea will be discussed in LVF 366-367.

Projection matrix

- Suppose $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. The projection of $\mathbf{u}$ to $\mathbf{v}$ is $\mathbf{v}(\mathbf{v}^T\mathbf{u})$.
- In $\mathbb{R}^n$, $\mathbf{v}^T\mathbf{u} = \mathbf{u}^T\mathbf{v} = \cos(\mathbf{u}, \mathbf{v})$.
- Matrix $\mathbf{P}_{(\mathbf{v})} = \mathbf{v}\mathbf{v}^T$ is called the projection matrix of subspace span$\{\mathbf{v}\}$.
- Vector $\mathbf{u}$ can be decomposed as $\mathbf{u} = \mathbf{u}_{(\mathbf{v})} + \mathbf{u}_{(\perp)}$, where

$$\mathbf{u}_{(\mathbf{v})} = \mathbf{v}(\mathbf{v}^T\mathbf{u}) = (\mathbf{v}\mathbf{v}^T)\mathbf{u} = \mathbf{P}_{(\mathbf{v})}\mathbf{u},$$

$$\mathbf{u}_{(\perp)} = \mathbf{u} - \mathbf{u}_{(\mathbf{v})} = (\mathbf{I} - \mathbf{v}\mathbf{v}^T)\mathbf{u}$$

- $\mathbf{P}_{(\perp)} = \mathbf{I} - \mathbf{v}\mathbf{v}^T$ is also a projection matrix, to the null subspace of span$\{\mathbf{v}\}$.

Gram-Schmidt using projection matrix

- Algorithm: Let $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \mathbf{a}_n]$

  1. Let $\mathbf{q}_1 = \mathbf{a}_1/\|\mathbf{a}_1\|$. $r_{11} = \|\mathbf{a}_1\|$
  2. For $i = 2, \cdots, n$

     (a) Project $\mathbf{a}_i$ to the null subspace of span$(\mathbf{q}_1, \cdots \mathbf{q}_{i-1})$,

     $$\begin{aligned} \mathbf{v}_i &= (\mathbf{I} - \mathbf{q}_1\mathbf{q}_1^T)\cdots(\mathbf{I} - \mathbf{q}_{i-1}\mathbf{q}_{i-1}^T)\mathbf{a}_i \\ &= (\mathbf{I} - \mathbf{Q}_{i-1}\mathbf{Q}_{i-1}^T)\mathbf{a}_i \ //\mathbf{Q}_{i-1} = (\mathbf{q}_1, \cdots \mathbf{q}_{i-1}) \end{aligned}$$

     (b) Normalization $\mathbf{q}_i = \mathbf{v}_i/\|\mathbf{v}_i\|$.

- Operation counts $\sim 2mn^2$

- Numerically unstable. $(\mathbf{Q}^T\mathbf{Q} = \mathbf{I}?\ \ \mathbf{QR} = \mathbf{A}?)$

Method II. Householder reflector LVF pp.161

- Like LU decomposition, Householder transformation zeros out elements $\mathbf{A}(i+1:n,i)$ column by column.

$$\begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{pmatrix} \Rightarrow \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{H}_1\mathbf{A} \Rightarrow \begin{pmatrix} 1 & \\ & \widehat{\mathbf{H}}_2 \end{pmatrix} \mathbf{H}_1\mathbf{A} \Rightarrow \begin{pmatrix} 1 & & \\ & 1 & \\ & & \widehat{\mathbf{H}}_3 \end{pmatrix} \mathbf{H}_2\mathbf{H}_1\mathbf{R} \Rightarrow \mathbf{H}_3\mathbf{H}_2\mathbf{H}_1\mathbf{A} = \mathbf{R}$$

- $\mathbf{Q} = (\mathbf{H}_3\mathbf{H}_2\mathbf{H}_1)^{-1} = \mathbf{H}_1^{-1}\mathbf{H}_2^{-1}\mathbf{H}_3^{-1} = \mathbf{H}_1^T\mathbf{H}_2^T\mathbf{H}_3^T$
  - $\mathbf{H}_i$ are orthogonal matrices. (So are $\widehat{\mathbf{H}}_i$)

Reflection matrix (reflector)

- Suppose a vector $\mathbf{w} = a\mathbf{u} + b\mathbf{v}$, where $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ and $\mathbf{u}^T\mathbf{v} = 0$. ($\mathbf{u}$, and $\mathbf{v}$ are orthogonal.)

- The reflection $\mathbf{w}'$ of $\mathbf{w}$ across $\mathbf{u}$ ($\mathbf{w}' + \mathbf{w} = \alpha\mathbf{u}$) is

$$\begin{aligned} \mathbf{w}' &= a\mathbf{u} - b\mathbf{v} = (a\mathbf{u} + b\mathbf{v}) - 2b\mathbf{v} \\ &= \mathbf{w} - 2(\mathbf{v}^T\mathbf{w})\mathbf{v} = (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)\mathbf{w} \end{aligned}$$

- The matrix $\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T$ is a reflection matrix (reflector).
  - $\mathbf{H}$ is symmetric: $\mathbf{H} = \mathbf{H}^T$.
  - $\mathbf{H}$ is orthogonal: $\mathbf{H}^{-1} = \mathbf{H}^T$.

QR decomposition using reflector LVF pp.170

- Design a reflector $\mathbf{H}$ s.t. for a given vector $\mathbf{x}$, $\mathbf{Hx} = \pm\|\mathbf{x}\|\mathbf{e}_1$.
  - $\mathbf{H}$ has the form $\mathbf{H} = \mathbf{I} - 2\frac{\mathbf{v}\mathbf{v}^T}{\|\mathbf{v}\|^2}$.
  - $\mathbf{v}$ is the *angle bisector* of $-\mathbf{e}_1$ and $\mathbf{x}$, $\mathbf{v} = \mathbf{x} \mp \|\mathbf{x}\|\mathbf{e}_1$.

Proof: Let $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$ and assume $\|\mathbf{v}\| = 1, \|\mathbf{x}\| = \alpha$.

$$\begin{aligned}
\mathbf{Hx} &= (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{x} - 2\mathbf{v}(\mathbf{v}^T\mathbf{x}) \\
&= \mathbf{x} - 2(\mathbf{v}^T\mathbf{x})(\mathbf{x} - \alpha\mathbf{e}_1) = (1 - 2\mathbf{v}^T\mathbf{x})\mathbf{x} + 2\alpha(\mathbf{v}^T\mathbf{x})\mathbf{e}_1 \\
1 - 2\mathbf{v}^T\mathbf{x} &= \mathbf{v}^T\mathbf{v} - 2\mathbf{v}^T\mathbf{x} = \mathbf{v}^T(\mathbf{v} - 2\mathbf{x}) \\
&= (\mathbf{x}^T - \alpha\mathbf{e}_1)(-\mathbf{x} - \alpha\mathbf{e}_1) = -\mathbf{x}^T\mathbf{x} + \alpha^2\mathbf{e}_1^T\mathbf{e}_1 = 0 \\
\mathbf{v}^T\mathbf{x} &= 1/2 \Rightarrow \mathbf{Hx} = \alpha\mathbf{e}_1 = \|\mathbf{x}\|\mathbf{e}_1
\end{aligned}$$

- Algorithm:

  1. Let $\mathbf{R}^{(0)} = \mathbf{A}$
  2. For $i = 1, \cdots, n-1$
     - Compute $\hat{\mathbf{H}}_i$ s.t. $\hat{\mathbf{H}}_i \mathbf{R}^{(i-1)}(i, i:n) = \|\mathbf{R}^{(i-1)}(i, i:n)\| \mathbf{e}_1$.
     - Let $\mathbf{H}_i = \begin{pmatrix} \mathbf{I} & \\ & \hat{\mathbf{H}}_i \end{pmatrix}$.
     - $\mathbf{R}^{(i)} = \mathbf{H}_i \mathbf{R}^{(i-1)}$
  3. $\mathbf{R} = \mathbf{R}^{(n-1)}$ and $\mathbf{Q} = \mathbf{H}_1^T \cdots \mathbf{H}_{n-1}^T$.

- Operation counts $\sim 2n^2(m - n/3)$ (assume $m > n$)
- Numerically stable.

Method III. Givens rotation LVF pp.168

- Find an orthogonal matrix $\mathbf{G}$ s.t. $\mathbf{G}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sqrt{a^2+b^2} \\ 0 \end{pmatrix}$

- Let $\begin{array}{l} \mathbf{G} = \begin{pmatrix} r_{11} & r_{12} \\ r_{22} & r_{22} \end{pmatrix} \\ r = \sqrt{a^2+b^2} \end{array} \Rightarrow \begin{cases} \gamma_{11}\gamma_{12} + \gamma_{21}\gamma_{22} = 0 & (1) \\ \gamma_{11}^2 + \gamma_{21}^2 = \gamma_{12}^2 + \gamma_{22}^2 = 1 & (2) \\ \gamma_{11}a + \gamma_{12}b = r & (3) \\ \gamma_{21}a + \gamma_{22}b = 0 & (4) \end{cases}$

  $r_{12}\times(3)+r_{22}\times(4) \Rightarrow (r_{11}r_{12}+r_{21}r_{22})a+(r_{12}^2+r_{22}^2)b=r_{12}r$
  $\Rightarrow r_{12} = \frac{b}{r}$. Plugin to (3), $\Rightarrow r_{11} = \frac{a}{r}$. Plugin to (2),(4)
  $\Rightarrow r_{21} = \frac{-b}{r}, r_{22} = \frac{a}{r}$.

- Define $\frac{a}{r} = \cos\theta, \frac{-b}{r} = \sin\theta, \Rightarrow \mathbf{G} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Plane rotation matrix

- In $\mathbb{R}^n$, the rotation matrix in the (i,j) plane is

$$
\mathbf{G}_{i,j,\theta} = \begin{array}{c c}
 & \begin{array}{ccccccccc} & & & i & & j & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i \\ \\ j \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & & & \\
 & \ddots & & & & & & & \\
 & & 1 & & & & & & \\
 & & & \cos\theta & & -\sin\theta & & & \\
 & & & & \ddots & & & & \\
 & & & \sin\theta & & \cos\theta & & & \\
 & & & & & & 1 & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & 1
\end{pmatrix}
\end{array}
$$

- $\mathbf{G}_{i,j,\theta}\mathbf{A}$ changes row i and row j of $\mathbf{A}$.
- $\mathbf{A}\mathbf{G}_{i,j,\theta}$ changes column i and column j of $\mathbf{A}$.

QR decomposition using rotation LVF pp.172

- Algorithm: zero out elements in the order

$$\begin{pmatrix} x & x & x & x & x \\ 4 & x & x & x & x \\ 3 & 7 & x & x & x \\ 2 & 6 & 9 & x & x \\ 1 & 5 & 8 & 10 & x \end{pmatrix}$$

- $\mathbf{Q} = \mathbf{G}_1^T \cdots \mathbf{G}_{n(n-1)/2}^T$, $\mathbf{R}$ is the remaining matrix.
- Operation counts $\sim 3n^2(m - n/3)$
- Numerically stable.

- Usually used on some special matrices, in which most elements are zeros
  - For example, upper Hessenberg matrix
  - A matrix $\mathbf{A}$ is upper Hessenberg if $a_{ij} = 0$ for $i > j + 1$.

$$\begin{pmatrix} x & x & x & x & x \\ 1 & x & x & x & x \\ & 2 & x & x & x \\ & & 3 & x & x \\ & & & 4 & x \end{pmatrix}$$

  - Operation count: $6n + 6(n-1) + \cdots + 6 \cdot 2 \approx 3n^2$.