LA2 8

KYB

Thrn, it's a Fact

mathrnfact@gmail.com

November 6, 2020

Overview

Operator Counts : the Augmented Matrix

Suppose A has a LU decomposition and A is nonsingular. We want to solve $Ax = b$. We can apply Gaussian elimination to the augmented matrix $[A|b]$.

- To make $[A|b] \to [U|c]$, $\frac{2}{3}n^3 + \frac{1}{2}n^2 - \frac{7}{6}n$ operations are required.
- To solve $Ux = b$, n^2 operations are required.

Hence the total number is $\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n$

Operator Counts : the LU factorization

Now instead of performing Gaussian elimination on the $[A|b]$, we operate on A to produce the factorization $A = LU$.

- To find L and U, $\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{1}{6}n$ operations are required.
- To solve $Lc = b$, $n^2 - n$ operations are required. (Ex 6.3.10)
- To solve $Ux = c$, n^2 operations are required.

Hence the total number is $\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n$.

That means the number of operations required by the two approaches is exactly the same.

An Advantage to using the LU factorization

In many problems, we do not just want to solve $Ax = b$ for a single vector b, but rather $Ax = b_i$ for k vectors $b_1, \cdots, b_k$. The expensive part of solving $Ax = b$ is Gaussian elimination or finding the LU factorization of A.
Since the matrix A is the same for every system, we need only compute the LU factorization once,

- To compute LU the factorization, $\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{1}{6}n \approx \frac{2}{3}n^3$ operations are required and ,
- To solve $LUx = b_i$, $k(2n^2 - n)$ operations are required.

Thus the total number is about $\frac{2}{3}n^3 + k(2n^2 - n)$.

On the other hand, if we were to perform Gaussian elimination and back substitution for each b_i, the total cost would be approximately $\frac{2k}{3}n^3$.

Matrix norms

Observation

Recall that a norm $\|\cdot\| : V \to \mathbb{R}$ is a function satisfying

- $\|v\| \geq 0$ with equality iff $v = 0$.
- $\|\alpha v\| = |\alpha|\|v\|$.
- $\|v + w\| \leq \|v\| + \|w\|$.

The set of $n \times n$ matrices $\mathbb{R}^{n\times n}$ is isomorphic to $\mathbb{R}^{n^2}$, a vector space over $\mathbb{R}$. Thus we can give a norm on $\mathbb{R}^{n\times n}$. On the other hand $\mathbb{R}^{n\times n}$ has an additional structure, the muliplication. Thus we may add a condition on a norm for matricses,

$$\|AB\| \leq \|A\|\|B\|.$$

More generally, if $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, add the same condition
If $\|\cdot\|$ is a norm satisfying the above condition, we call it a *matrix norm*.
Roughly speaking, a matrix norm is a function measuring the size of an operator.

Example

Give a norm $\|\cdot\|$ on $\mathbb{R}^{m\times n}$ by the Euclidean 2-norm as $\mathbb{R}^{mn}$. We call this norm the Frobenius norm. The Frobenius norm is a matrix norm.
Let $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$. Then

$$\|Ax\|_2^2 = \|Ax\|_2^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right)^2 \leq \sum_{i=1}^{m}\left(\left(\sum_{j=1}^{n} A_{ij}^2\right)\left(\sum_{j=1}^{n} x_j^2\right)\right)$$

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}\right)^2 \|x\|_2^2 = \|A\|_F^2\|x\|_2^2.$$

Example

In general,

$$\|A\|_F^2 = \sum_{i=1}^{m} \|A_i\|_2^2.$$

Hence

$$\begin{aligned}\|AB\|_F^2 &= \sum_{i=1}^{m} \|AB_i\|_2^2 \le \sum_{i=1}^{m} \|A\|_F^2 \|B_i\|_2^2 \\ &= \|A\|_F^2 \sum_{i=1}^{m} \|B_i\|_2^2 = \|A\|_F^2 \|B\|_F^2.\end{aligned}$$

Lemma (391)

If $A \in \mathbb{R}^{m \times n}$ *and* $x \in \mathbb{R}^n$, *then*

$$\|Ax\|_2 \leq \|A\|_F \|x\|_2.$$

Theorem (392)

For any $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$,

$$\|AB\|_F \leq \|A\|_F \|B\|_F.$$

We can apply the ϵ-δ definition of the continuity on an operator $x \mapsto Ax$. In this sence, A is continuous. Moreover, A is differentible and $(Ax)' = A$.

Definition

Let $\|\cdot\|$ represent a family of norms on $\mathbb{R}^n$. The induced matrix norm on $\mathbb{R}^{m\times n}$ is defined by

$$\|A\| = \sup\{\|Ax\| \ : \ x \in \mathbb{R}^n, \|x\| = 1\}.$$

Theorem (395)

Let $\|\cdot\|$ be a vector norm on $\mathbb{R}^n$. Then the induced matrix norm $\|\cdot\|$ is a norm on $\mathbb{R}^{n\times n}$.

Theorem (396)

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and let the induced matrix norm on $\mathbb{R}^{m\times n}$ be denoted by the same symbol. Then, for all $A \in \mathbb{R}^{m\times n}$,

$$\|A\| = \sup\left\{\frac{\|Ax\|}{\|x\|} \mid x \in \mathbb{R}^n, x \neq 0\right\}.$$

Corollary (397)

Let $\|\cdot\|$ *be a norm on* $\mathbb{R}^n$ *and let the same symbol denote the induced matrix norm on* $\mathbb{R}^{n\times n}$*. Then,*

$$\|Ax\| \leq \|A\|\|x\|.$$

Theorem (398)

Let $\|\cdot\|$ denote the matrix norm on $\mathbb{R}^{m\times n}$ induced by a vector norm $\|\cdot\|$. Then

$$\|A\| = \inf\{M > 0 \ : \ \|Ax\| \le M\|x\| \text{ for all } x \in \mathbb{R}^n\}.$$

Corollary (399)

Let $\|\cdot\|$ be a matrix norm induced by a vector norm $\|\cdot\|$. Then

$$\|AB\| \le \|A\|\|B\|.$$

Theorem (400)

The matrix norm induced by the l^1 vector norm satisfies

$$\|A\|_1 = \max \left\{ \sum_{i=1}^{m} |A_{ij}| \mid j = 1, 2, \cdots, n \right\}.$$

Theorem (401)

The matrix norm induced by the l^∞ vector norm satisfies

$$\|A\|_\infty = \max\left\{\sum_{j=1}^{n} |A_{ij}| \mid j = 1, 2, \cdots, m\right\}.$$

Theorem (402)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and let the eigenvalues of A be $\lambda_n \leq \lambda_{n-1} \leq \cdots \leq \lambda_1$. Then

$$\lambda_n \|x\|_2^2 \leq x \cdot (Ax) \leq \lambda_1 \|x\|_2^2.$$

Proof

Let $\{x_1, \cdots, x_n\}$ be an orthonormal sets such that $Ax_i = \lambda_i x_i$. Then every $x \in \mathbb{R}^n$ is of the form

$$x = \sum (x \cdot x_i) x_i.$$

(Continued)

Proof

$$
\begin{aligned}
x \cdot Ax &= \left(\sum (x \cdot x_i) x_i\right) \cdot A \left(\sum (x \cdot x_j) x_j\right) \\
&= \sum_i \sum_j (x \cdot x_i)(x \cdot x_j) x_i \cdot A x_j \\
&= \sum_i \sum_j (x \cdot x_i)(x \cdot x_j) x_i \cdot \lambda_j x_j = \sum_i (x \cdot x_i)^2 \lambda_i .
\end{aligned}
$$

Note that $\|x\|^2 = \sum (x \cdot x_i)^2$.

$$
\begin{aligned}
\lambda_n \sum_i (x \cdot x_i)^2 \leq & x \cdot Ax \leq \lambda_1 \sum_i (x \cdot x_i)^2, \\
\lambda_n \|x\|^2 \leq & x \cdot Ax \leq \lambda_1 \|x\|^2 .
\end{aligned}
$$

Theorem (403)

Let $A \in \mathbb{R}^{m\times n}$ *and let* λ_1 *be the largest eigenvalue of* $A^T A$*. Then* $\|A\|_2 = \sqrt{\lambda_1}$*.*

Corollary (404)

$\|A\|_2 = \sigma_1$ *where* σ_1 *is the largest singular value of* A*.*

Exercises

Let $A, B \in \mathbb{R}^{n\times n}$ such that A is invertible and $\|A-B\|\left\|A^{-1}\right\| < 1$, where $\|\cdot\|$ is a induced matrix norm from $\|\cdot\|_2$. Prove that B is invertible.

Proof

Suppose $Bx = 0$.

$$\begin{aligned}\|x\|_2 = \left\|A^{-1}Ax\right\|_2 &\leq \left\|A^{-1}\right\|\|Ax\|_2 = \left\|A^{-1}\right\|\|(A-B)x\|_2 \\ &\leq \left\|A^{-1}\right\|\|A-B\|\|x\|_2\end{aligned}$$

If $\|x\|_2 \neq 0$, $\|x\|_2 \leq \left\|A^{-1}\right\|\|A-B\|\|x\|_2 < \|x\|_2$. So $0 < \|x\|_2 < \|x\|_2$ (contradiction). Hence $\|x\|_2 = 0$, or $x = 0$.

Remark

Take $A = I$. For all $\|x\|_2 = 1$, $\|Ix\|_2 = \|x\|_2 = 1$. So $\|I\| = 1$. Then if $\|B-I\| < 1$, B is invertible.

Caution

- $\|I\|_F = \sqrt{\sum_{i=1}^{n} 1} = \sqrt{n}$.
- We know that if A is invertible, $\det(A^{-1}) = \det(A)^{-1}$. But $\|A^{-1}\| \neq \|A\|^{-1}$ in general.

For example,

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\|A\|_2 = 2, \|A^{-1}\|_2 = 1,$$
$$\|A\|_F = \sqrt{5}, \|A^{-1}\|_F = \sqrt{5}/2$$

Remark

(Ch 9.5) Given induced matrix norm $\|\cdot\|$, we can define the *condition number* for invertible matrix A by

$$\operatorname{cond}(A) =: \|A\| \left\|A^{-1}\right\|.$$

For l^2 norm,

$$\operatorname{cond}_2(A) = \|A\|_2 \left\|A^{-1}\right\|_2 = \frac{\sigma_1}{\sigma_n}.$$

If A is singular, define $\operatorname{cond}(A) = \infty$.

Ex 9.4.1

Let $\|\cdot\|$ be any induced matrix norm on $\mathbb{R}^{n\times n}$. Prove that $\rho(A) \leq \|A\|$ for all $A \in \mathbb{R}^{n\times n}$, where $\rho(A)$ is the spectral radius of A:

$$\rho(A) = \max\{|\lambda| \ : \ \lambda \text{ is an eigenvalue of } A\}.$$

Ex 9.4.7

Let $A \in \mathbb{R}^{n \times n}$. Prove that $\left\|A^T\right\|_2 = \|A\|_2$.

The QR factorization

Observation

Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$, of full rank. Then $\{A_1, \cdots, A_n\}$ is a basis of $\mathrm{col}(A)$. Now by apllying the Gram-Schmidt process,

$$q_1 = A_1, Q_1 = \frac{q_1}{\|q_1\|_2},$$

$$\vdots$$

$$q_n = A_n - \sum_{i=1}^{n-1} (A_n \cdot Q_i) Q_i, Q_n = \frac{q_n}{\|Q_n\|_2}.$$

Observation

As a linear system, the Gram-Schmidt process can be written by

$$\begin{aligned} A_1 &= r_{11}Q_1 \\ A_2 &= r_{12}Q_1 + r_{22}Q_2 \\ \vdots \quad &\quad \vdots \\ A_n &= r_{1n}Q_1 + r_{2n}Q_2 + \cdots + r_{nn}Q_n \end{aligned}$$

So,

$$A = \underbrace{\left[Q_1 \mid Q_2 \mid \cdots \mid Q_n\right]}_{Q_0} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}}_{R_0}$$

Then $Q_0 \in \mathbb{R}^{m\times n}$ and $R_0 \in \mathbb{R}^{n\times n}$. Moreover $\det(R_0) \neq 0$.

Observation

Let R_0 be an upper triangular matrix whose entries $(R_0)_{ij} = r_{ij}$ and $Q_0 = [Q_1|\cdots|Q_n]$, and then we get $A = Q_0R_0$. Now extend $\{Q_1, \cdots, Q_n\}$ to an orthonormal basis $\{Q_1, \cdots, Q_m\}$ and add an $(m-n) \times n$ block of zeros to the end of the matrix R_0 to product an $m \times n$ upper triangular matrix R.

$$QR = \left[Q_0 \mid \tilde{Q}_0\right] \left[\frac{R_0}{0}\right] Q_0R_0 + \tilde{Q}_0 0 = Q_0R_0 = A.$$

Application

Given the QR factorization of $A \in \mathbb{R}m \times n$,

$$A^TA = (Q_0R_0)^T(Q_0R_0) = R_0^TQ_0^TQ_0R_0 = R_0^TR_0$$

$$A^TAx = A^Tb \Leftrightarrow R_0^TR_0x = R_0^TQ_0^Tb \Leftrightarrow R_0x = Q_0^Tb.$$

Householder transformation

Suppose A has a QR factorization. Then $Q^T A = R$ implies $Q^T : A_1 \mapsto r_{11} e_1$, where $r_{11} = \|A_1\|_2$.

Let $v = A_1$. If U is an orthogonal matrix such that $Uv = \alpha e_1$. Since U preserves a norm, $\|Uv\|_2 = \|v\|_2 = |\alpha|$. So $\alpha = \pm\|v\|_2$.

Householder transformation

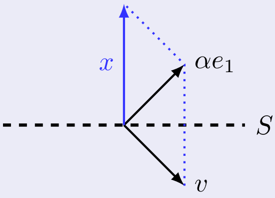
Let $x = \alpha e_1 - v$ where $\alpha = \pm\|v\|_2$ and $S = x^{\perp}$. Define $u_1 = x/\|x\|_2$ and let $\{u_2, \cdots, u_m\}$ be an orthonormal basis for S.

$$U = \left[u_1 \,\middle|\, \cdots \,\middle|\, u_m\right] = I - 2u_1 \otimes u_1$$

is an orthogonal matrix and by Ex 9.8.7 there exists $y \in S$ such that

$$v = y - \frac{1}{2}x, \alpha e_1 = y + \frac{1}{2}x.$$

Then $Uy = y$ and $Uv = \alpha e_1$, and U defines a reflection across S.



Ex. 9.8.7

Let $x \in \mathbb{R}^m$ be given, define $\alpha = \pm\|v\|_2$, $x = \alpha e_1 - v$, $u_1 = x/\|x\|_2$, let $\{u_1, u_2, \cdots, u_m\}$ be an orthonormal basis for $\mathbb{R}^m$, and let

$$S = \{u_1\}^{\perp} = \mathrm{span}\{u_2, \cdots, u_m\}.$$

Prove that there exists $y \in S$ such that

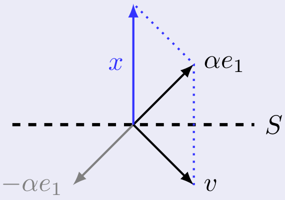
$$v = y - \frac{1}{2}x, \alpha e_1 = y + \frac{1}{2}x.$$

QR factorization using Householder transformation

Let $v = A_1$ and choose u such that $u = x/\|x\|$, $x = \alpha_1 e_1 - v$, $\alpha_1 = -\operatorname{sgn}(v_1)\|v\|_2$. Define $Q_1 = I - 2uu^T$ and compute $A^{(2)} = Q_1 A$. Then

$$A^{(2)} = \left[\begin{array}{c|c} \alpha_1 & a_1^T \\ \hline 0 & B^{(2)} \end{array}\right]$$

The reason that we take $-\operatorname{sgn}(v_1)$ is just to avoid the subtracting error (numerical reason). Theorically, the sign does not matter.



QR factorization using Householder transformation

Suppose $Q_1, \cdots, Q_k$ are given. Then $A^{(k+1)} = Q_k Q_{k-1} \cdots Q_1 A$.

$$Q_k A^{(k)} = A^{(k+1)} = \left[\begin{array}{c|c} R^{(k+1)} & C^{(k+1)} \\ \hline 0 & B^{(k+1)} \end{array}\right]$$

Find $\tilde{Q}_{k+1} \in \mathbb{R}^{(n-k)\times(n-k)}$ so that

$$\tilde{Q}_{k+1} B^{(k+1)} = \left[\begin{array}{c|c} \alpha_k & a_k^T \\ \hline 0 & B^{(k+2)} \end{array}\right]$$

Define

$$Q_{k+1} = \left[\begin{array}{c|c} I_{n-k} & 0 \\ \hline 0 & \tilde{Q}_{k+1} \end{array}\right].$$

QR factorization using Householder transformation

Indeuctively, $A^{(n+1)} = Q_n Q_{n-1} \cdots Q_1 A = \left[\dfrac{R_0}{0}\right] = R$, where $R_0 \in \mathbb{R}^{n\times n}$ is upper triangular with nonzero diagonal entries. Then

$$A = (Q_n Q_{n-1} \cdots Q_1)^T R = QR$$

Ex 9.8.1

Find a Householder transformation $Q = I - 2uu^T$ satisfying $Qx = y$ where $x = (1, 2, 1)$ and $y = (2, 1, 1)$.

Ex 9.8.6

Let $\{u_1, \cdots, u_m\}$ be an orthonormal basis for $\mathbb{R}^m$, let

$$X = [u_1 | \cdots | u_m]$$

and let $D \in \mathbb{R}^{m \times m}$ be the diagonal matrix whose diagonal entries are $-1, 1, \cdots, 1$. Finally, define $U = XDX^T$.

(a) Prove that $U = -u_1 u_1^T + u_2 u_2^T + \cdots + u_m u_m^T$.

(b) Show that $u_1 u_1^T + u_2 u_2^T + \cdots + u_m u_m^T = I$.

(c) Show that the formular for U simplifies to $U = I - 2u_1 u_1^T$.

(d) Verify directly that U is symmetric and orthogonal.

The End