Math 571 - Numerical Linear Algebra

Autumn 2024

August 29, 2024

27 Aug 2024

There are three types of problems that arise in numerical linear algebra.

1. Suppose $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^{n\times 1}$ is given and A is invertible. Then solve $Ax = b$.
2. Given $A \in \mathbb{R}^{n\times m}$ with $n > m$ and $b \in \mathbb{R}^n$, solve $Ax = b$.
3. $A \in \mathbb{C}^{n\times n}$ is given. Find $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n$, $x \neq 0$, such that $Ax = \lambda x$.

There are direct and iterative methods. Sometimes A is large, but sparse, and iterative methods make a little more sense. Problems of type 3 must be iterative by nature.

Example. **C**onsider the temperature along a wire. $0 \le x \le 1$ with varying insulation thickness given by $u(x,t)$. We have $u_t = u_{xx} - f(x)$, with the initial condition $u(x,0) = u_0(x)$. Furthermore, $u(0,t) = \alpha$ and $u(1,t) = \beta$. Is there a steady state temperature distribution along the wire?

Consider the steady state $v(x)$ which must solve $v''(x) = f(x)$ and $v(0) = \alpha, v(1) = \beta$. To approximate, we divide the interval $[0,1]$ into $N+1$ units (grid spacing $h = 1/(N+1)$) and solve the system $v_j = v(x_j)$. We need to deduce the unknowns v_1, v_2, ..., v_N. For $j \in \{2,3,\ldots,N-1\}$, we can use the approximation

$$v''(x_j) \approx \frac{v_{j+1} - 2v_j + v_{j-1}}{h^2}.$$

Furthermore, for $j = 1$, we can note that

$$v''(x_1) \approx \frac{v_2 - 2v_1 + \alpha}{h^2} = f_1$$

and for $j = N$, we have

$$v''(x_N) \approx \frac{\beta - 2v_N + v_{N-1}}{h^2} = f_N.$$

This is a linear system of N equations in N unknowns $v_1, \ldots, v_N$. This is alinear system of N equations in N unknowns $v_1, \ldots, v_N$:

$$\begin{pmatrix} -2 & 1 & 0 & \ldots & 0 \\ 1 & -2 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 1 & -2 & 1 \\ 0 & \ldots & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{N-1} \\ v_N \end{pmatrix} = \begin{pmatrix} h^2 f_1 - \alpha \\ h^2 f_2 \\ \vdots \\ h^2 f_{N-1} \\ h^2 f_N - \beta \end{pmatrix}$$

This matrix is very sparse (fewer than $3N$ nonzero entries) which are clustered along the diagonal. This is known as a *tridiagonal* matrix. This system can be solved for given RHS in $O(n)$ operations using Gaussian elimination.

Fact: in the previous example, M_N^{-1} is not sparse at all. In fact, it is almost completely dense. Consider an $N \times N$ matrix A and multiply by an $N \times 1$ vector w. Then, computing Aw involves $O(N^2)$ operations, so we will basically never form A^{-1} when solving $Ax = b$.

We will look at Gaussian elimination, QR factorization, etc as direct methods, which we will interpret as matrix factorization.

A few simple facts:

- $A \in \mathbb{R}^{n \times m}$, $A : \mathbb{R}^m \mapsto \mathbb{R}^n$.
- $(Ax)_i = \sum_j A_{ij} x_j$.
- If A_i refers to the i=th column of a matrix, then

$$Ax = x_1 A_1 + x_2 A_2 + \ldots + x_m A_m$$

Definition. A *norm* on $\mathbb{C}^n$ is a function $|\cdot| : \mathbb{C}^n \to \mathbb{R}^+$ satisfying:
1. $|x| \geq 0$ and $|x| = 0$ iff $x = 0$.
2. For any $c \in \mathbb{C}$ and $x \in \mathbb{C}^n$, $|cx| = |c||x|$.
3. $|x + y| \leq |x| + |y|$.

Definition. We can consider the *Euclidean norm* $|x|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$. Furthermore, the *p-norm* is defined by $|x|_p = (|x_1|^p + \ldots + |x_n|^p)^{1/p}$ for $1 \leq p \leq \infty$ ($p = 1$ is useful). Finally, $|x|_\infty = \max_{j=1,\ldots,n} |x_j|$.

Definition. Take $A \in \mathbb{R}^{n \times m}$. Then $||A||_F = \sqrt{\sum_{i,j} A_{i,j}^2}$. This is the *Frobenius (Hilbert-Schmidt) norm.*

Definition. If $A \in \mathbb{R}^{n \times m}$, we can consider the *induced p-norm* defined by considering

$$\max_{x \neq 0} \frac{|Ax|_p}{|x|_p} = \max_{|x|=1} |Ax|_p.$$

One can note that $p = 1$ and $p = \infty$ are easier to work with than $p = 2$. First, let us consider the induced $p = 1$ norm of a matrix. Note that

$$\begin{aligned}
|Ax|_1 &= |A_1 x_1 + A_2 x_2 + \ldots + A_m x_m| \\
&\leq |A_1 x_1| + |A_2 x_2| + \ldots + |A_m x_m| \\
&= |x_1||A_1| + |x_2||A_2| + \ldots + |x_m||A_m| \\
&\leq |a_1||A_{i_*}| + |a_2||A_{i_*}| + \ldots + |a_n||A_{i_*}|
\end{aligned}$$

where $|A_{i_*}| = \max_{1 \leq j \leq m} |A_j|$. Thus,

$$|Ax| \leq \underbrace{(|x_1| + \ldots + |x_m|)}_{\text{sum}=1} |A_{i_*}| = |A_{i_*}|$$

so we get the bound $|A|_1 \leq |A_{i_*}|$. It is easy to see the upper bound is attained by chosing an appropriate standard basis vector, so it follows that $|A|_1 \leq |A_{i_*}|$.

29 August 2024

First, we can talk a little more about the Frobenius norm on $A \in \mathbb{C}^{n\times m}$. We can note that

$$\left(\sum_{i,j} |A_{i,j}|^2\right)^{1/2} = \sqrt{\sum_{i,j} \overline{A_{i,j}} A_{i,j}} = \sqrt{\sum_{i,j} (A^*)_{j,i} A_{i,j}} = \sqrt{\sum_{j} (A^*A)_{j,j}} = \sqrt{\mathrm{tr}(A^*A)}.$$

Furthermore, we discussed the *induced norm* defined on $1 \le p \le \infty$ defined by

$$||A||_p = \max_{|x|_p=1} |Ax|_p = \max_{x\neq 0}.$$

Recall $||\cdot||_1$ and $||\cdot||_\infty$ are often the easiest induced matrix norms to compute. We have already seen how to compute the $||\cdot||_1$ norm so now we will attempt to compute $||\cdot||_\infty$. We have

$$|(Ax)_i| = |\sum_j A_{i,j}x_j| \le \sum_j |A_{i,j}||x_j| \le \sum_j |A_{i,j}|.$$

This is equal to the absolute row sum along the row i. Thus,

$$|Ax|_\infty = \max_i |(Ax)_i| \le \max_i \sum_j |A_i, j|.$$

It is easy to find a value of x such that the inequality turns into an equality.

Finally, we can talk about the induced 2-norm on matrices. We have the following result:

Theorem. If $A \in \mathbb{C}^{n\times m}$, then $||A||_2 = \sqrt{\lambda}$, where λ is the largest eigenvalue of A^*A.

Proof. Suppose $A \in \mathbb{C}^{n\times m}$. Then, by definition,

$$||A||_2^2 = \max_{|x|_2=1} |Ax|_2^2,$$

and we can write $|Ax|_2^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle$ under the usual inner product. Thus, our optimization problem turns into

$$||A||_2^2 = \max_{|x|_2=1} \langle A^*Ax, x\rangle$$

where $A^*A \in \mathbb{C}^{m\times M}$. Now, A^*A is Hermitian, so A^*A is *unitarily diagonizable* (eigenvectors can be chosen to form an orthonormal basis). Moreover, all eigenvalues are real and positive. Take $\{v_1, \dots, v_m\}$ to be the orthonormal basis satisfying $A^*Av_j = \lambda j v_j$. If we take $V = (v_1^T, \dots, v_n^T)$ and

$$\Sigma = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{pmatrix},$$

then $A^*A = V\Sigma V^*$. Thus,

$$|Ax|_2^2 = \max_{|x|_2=1} \langle A^*Ax, x\rangle = \max_{|x|_2=1} \langle \Sigma V^*x, V^*x\rangle.$$

Since $V^* = V^{-1}$, it follows that

$$\langle V^*x, V^*x\rangle = \langle x, VV^*x\rangle = \langle x, x\rangle,$$

so if we set $y = V^*x$, it follows that

$$\begin{aligned}
\max_{|x|_2=1} \langle \Sigma V^* x, V^* x \rangle &= \max_{|y_1|^2+\ldots+|y_m|^2=1} \langle \Sigma_y, y \rangle \\
&= \max_{|y_1|^2+\ldots+|y_m|^2} \left(\lambda_1 |y_1|^2 + \lambda_2 |y_2|^2 + \ldots + \lambda_m |y_m|^2 \right) \\
&= \max_j \lambda_j .
\end{aligned}$$

From this, we can conclude that $||A||_2^2$ is the largest eigenvalue of A^*A. □

Suppose $A \in \mathbb{C}^{m\times m}$ and suppose that A is diagonalizible. Then, $A = MDM^{-1}$ for some matrix M and diagonal matrix D. If $A = A^*$, then A is unitarily diagonalizible, so we can write $A = UDU^*$ for some unitary matrix U. Singular value decomposition generalizes this to all matrices $A \in \mathbb{C}^{n\times m}$.

Definition. If $A \in \mathbb{C}^{n\times m}$, the *Singular Value Decomposition (SVD)* of A is of the form

$$A = U\Sigma V^*$$

where V has dimensions $m \times m$ and unitary, Σ is diagonal and has dimensions $ntimesm$, and U has dimensions $N \times n$ and is unitary.

Assuming SVD exists, we can find the SVD by noting first that

$$A^*A = M_1 D_1 M_1^* \tag{1}$$

where M_1 is unitary, D_1 is diagonal and has positive entries, and all matrices have dimensions $m \times m$. Furthemore,

$$AA^* = M_2 D_2 M_2^{-1}, \tag{2}$$

where M_2 is unitary and D_2 is diagonal and positive. By substituting $A = U\Sigma V^*$ into (1), we can conclude that

$$A^*A = V\Sigma^T U^* U \Sigma V^* = V\Sigma^T \Sigma V^* = M_1 D_1 M_1^*$$

Thus, if Σ is diagonal with entries $\sigma_1, \ldots, \sigma_k$ with $K = \min n, m$, then $\sigma_j = \sqrt{\lambda_j}$ where λ_j are eigenvaleus of A^*A. Furthermore, columns of V are eigenvectors $v_1, \ldots, v_m$ of A^*A. To find U, we note that

$$Av_j = U\Sigma V^{-1} v_j = U\Sigma e_j = U\Sigma_j$$

where e_j is the j-th standard basis vector and Σ_j is the j-th column. Since $U\Sigma_j = \sigma_j U e_j = \sigma_j U_j$, it follows that $Av_j = \sigma_j U_j$. This completes the algorithm to find the SVD of a matrix A.