# Numerical Analysis 1

Tyler Chang
February 21, 2018

## Stability

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Let $f(x) = y$ and let $\hat{f}(x) = \hat{y}$ be the computed value of $f(x)$, and assume that for some $\hat{x} \in \mathbb{R}^n$, $f(\hat{x}) = \hat{y}$.

- Forward error: $\|y - \hat{y}\|$
- Backward error: $\|x - \hat{x}\|$
- Condition number is the smallest $M$ such that for all $x$, $\frac{\|y-\hat{y}\|}{\|y\|} \leq M \frac{\|x-\hat{x}\|}{\|x\|}$

As a linear operator, a matrix has a condition number. For a norm $\|.\|$, the condition number of $A$ is given by $cond(A) = \|A^{-1}\|\|A\|$ and can be interpreted as closeness to singularity.

## Linear Solves

Solve $\mathbf{Ax} = \mathbf{b}$ using gaussian elimination. This can also be expressed as finding the LU decomposition $\mathbf{A} = \mathbf{LU}$ and solving

$$\mathbf{Ux} = \mathbf{L}^{-1}\mathbf{b}$$

where $\mathbf{L}$ and $\mathbf{U}$ are lower and upper triangular respectively.

If $\mathbf{A}$ is poorly conditioned, we may need to swap some of its rows to ensure stability. These row swaps can be captured by adding a permutation matrix $\mathbf{P}$:

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{LU}.$$

Note that $\mathbf{P}$ will be the identity matrix with its columns shuffled. If $\mathbf{A}$ is sparse, we can instead pivot to avoid fill-in.

If $\mathbf{A}$ is SPD, then there exists a nonsingular lower triangular matrix $\mathbf{L}$ and diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{LDL}^T.$$

This is called the Cholesky decomposition and is faster than the LU decomposition. Then we can solve $\mathbf{Ax} = \mathbf{b}$ by

$$\mathbf{DL}^T\mathbf{x} = \mathbf{L}^{-1}\mathbf{b}.$$

## Least Squares

Find solution $\mathbf{x}^*$ to minimize $\|\mathbf{Ax} - \mathbf{b}\|_2$ over $\mathbf{x}$. Note that a necessarry and sufficient condition is that the residual vector $\mathbf{b} - \mathbf{Ax}^*$ must be orthogonal to the range of $\mathbf{A}$. It follows that $\mathbf{x}^*$ is the projection of $\mathbf{b}$ onto the range of $\mathbf{A}$.

One solution (Normal Equations) is to solve:

$$\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}.$$

Another solution (Gram-Schmidt) is to compute an orthonormal basis for the range of $\mathbf{A}$ and use it to locate the projection of $\mathbf{b}$ then take this solution as $\mathbf{x}^*$. This is equivalent to decomposing $\mathbf{A}$ into an orthonormal matrix $\mathbf{Q}$ and an upper triangular $\mathbf{R}$ such that $\mathbf{A} = \mathbf{QR}$ then solving $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$, for all the nonzero rows of $\mathbf{R}$. The $\mathbf{QR}$ decomposition can also be done with householder reflectors (which is computationally equivalent to Gram-Schmidt) or Givens rotations.

If $\mathbf{A}$ is rank-deficient, then there are infinitely many solutions. To obtain a minimum norm solution, take the SVD of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

and solve

$$\mathbf{x}^* = \mathbf{V}_r\mathbf{\Sigma}_r^{-1}\mathbf{U}^T\mathbf{b},$$

where $\mathbf{V}_r$ and $\mathbf{\Sigma}_r$ denote the first $r$ columns of $\mathbf{V}$ and rows of $\mathbf{\Sigma}$ respectively, and where $r$ is the rank of $\mathbf{\Sigma}$.

## Eigenvalue Problems

Find all eigenpairs of $\mathbf{A}$: $(\lambda, x)$ satisfying $\mathbf{Ax} = \lambda\mathbf{x}$. Note that any matrix $\mathbf{B}$ such that $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ is similar to $\mathbf{A}$ in that it has the same eigenpairs as $\mathbf{A}$. We will say that $\mathbf{A}$ is diagonalizable if it is similar to a diagonal matrix. Note that if we know the eigenvalues of $\mathbf{A}$, we can find the eigenvectors by solving a linear system.

**Schur Decomposition**:
$\mathbf{U}^*\mathbf{AU} = \mathbf{T}$ where $\mathbf{U}$ is unitary, $\mathbf{U}^*$ denotes conjugate transpose, and $\mathbf{T}$ is upper triangular with the eigenvalues of $\mathbf{A}$ on its diagonal.

**Gershigoran Disk Thm**:
If $\mathbf{A} \in \mathbb{C}^{n\times n}$, then there are $n$ disks in the complex plane centered at each of the diagonals of $\mathbf{A}$ with radius equal to the sum of the nondiagonal elements in each column. All the eigenvalues of $\mathbf{A}$ are contained in these disks.

Using the fact that $\mathbf{A}$ and $\mathbf{S}^*\mathbf{AS}$ have the same number of positive/negative/zero eigenvalues, we can use bisection to search for them by scaling a diagonal matrix $\mathbf{S}$.

We can also use vector iterations: pick a vector $\mathbf{x}$ and multiply by $\mathbf{A}$ until it converges to something (the largest eigenvalued eigenvector).

Also see subspace iteration, QR iteration, inverse iteration, Rayleigh quotient iteration.

## Iterative Methods

Again, we want to solve $\mathbf{Ax} = \mathbf{b}$, but this time we want to solve with an iterative method. Split $\mathbf{A}$ into $\mathbf{A} = \mathbf{M} - \mathbf{N}$ where $\mathbf{M}$ is something easily invertible (to get the Jacobi iterative, just take $\mathbf{M}$ to be the diagonal of $\mathbf{A}$). Then $\mathbf{x} = \mathbf{M}^{-1}\mathbf{N}\mathbf{x} + \mathbf{M}^{-1}\mathbf{b}$. Iterate by

$$\mathbf{x}_{k+1} = \mathbf{M}^{-1}\mathbf{N}\mathbf{x}_k + \mathbf{M}^{-1}\mathbf{b}.$$

One can also perform overrelaxation by over stepping in the direction of each iterate by a constant factor of $\tau > 1$. In some cases, this may accelerate convergence since $x_k \to x_{k+1}$ may advance slowly. One can use vector accelleration by writing the step as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{M}^{-1}\mathbf{r}_k$$

where $\mathbf{r_k} = \mathbf{b} - \mathbf{A}\mathbf{x}_k$, then modifying to

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tau_k\mathbf{M}^{-1}\mathbf{r}_k$$

where $\tau_k$ is chosen variably at each step to maximize convergence by some metric.

Note that the above accellerated methods always produce solutions $\mathbf{x}_k \in span\{\mathbf{b}, \hat{\mathbf{A}}\mathbf{b}, \hat{\mathbf{A}}^2\mathbf{b}, \ldots\}$ where $\hat{\mathbf{A}} = \mathbf{M}^{-1}\mathbf{A}$. This is called the Krylov space, and Krylov methods reduce complexity by restricting the search for $\mathbf{x}$ to the Krylov space: $span\{\mathbf{b}, \hat{\mathbf{A}}\mathbf{b}, \hat{\mathbf{A}}^2\mathbf{b}, \ldots\}$. The Arnoldi iteration is one such method and is used to solve least squares problems in `Lapack GMRES`.