CHAPTER 9.3 Iterative Methods

Iterative Methods of Solving Linear Systems

Roundoff error can actually accelerate the convergence of an iteraitive method towards a solution

JACOBI'S METHOD

$$
\begin{aligned}
7x_1 - x_2 &= 5 \\
3x_1 - 5x_2 &= -7
\end{aligned}
$$

Solve the first equation for x_1 and the second equation for x_2

$$
\begin{aligned}
x_1 &= \frac{5 + x_2}{7} \\
x_2 &= \frac{7 + 3x_1}{5}
\end{aligned}
$$

Then use $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ as an initial approximation.

$$
\begin{aligned}
x_1 &= \frac{5+0}{7} = \frac{5}{7} \approx 0.714 \\
x_2 &= \frac{7+3\ 0}{5} = \frac{7}{5} = 1.400
\end{aligned}
$$

Substitute these values into the second iteration

$$
\begin{aligned}
x_1 &= \frac{5+1.4}{7} \approx 0.914 \\
x_2 &= \frac{7+3\ \frac{5}{7}}{5} \approx 1.829
\end{aligned}
$$

After several iterations we can see a convergence to $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

n	0	1	2	3	4	5	6
x_1	0	0.714	0.914	0.976	0.993	0.998	0.999
x_2	0	1.400	1.829	1.949	1.985	1.996	1.999

GAUSS-SEIDEL METHOD

The method is the same as Jacobi's except that we use each new value as soon as we can. In other words, calculate x_1 as we did before, but then use the result ($\frac{5}{7}$ in this case) to get the next value of x_2.

$$x_2 = \frac{7 + 3\,\frac{5}{7}}{5} \approx 1.829$$

This provides a quicker convergence

n	0	1	2	3	4	5
x_1	0	0.714	0.976	0.998	1.000	1.000
x_2	0	1.829	1.985	1.999	2.000	2.000

ITERATIVE METHODS FOR EIGENVALUES

The Power Method

This is for a matrix with a DOMINANT EIGENVALUE: one that is larger in absolute values than all of the other eigenvalues.

The power method proceeds iteratively to produce a sequence of scalars that converge to the dominant eigenvalue and a sequence of vectors that converges to the corresponding eigenvector.

For simplicity, let A be diagonalizable with a dominant eigenvalue. There exists a nonzero vector x_0 such that the sequence of vectors x_k defined by

$$x_1 = Ax_0; x_2 = Ax_1; x_k = Ax_{k-1} \ldots$$

approaches the dominant eigenvector of A.

Proof:

Eigenvalues of A would be listed as $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Corresponding eigenvectors would be independent (A is diagonalizable) and would correspond $v_1, v_2, \cdots, v_n$. Since there are a basis they can be expressed as a linear combination of eigenvectors: $x_0 = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$

Now $x_1 = Ax_0, x_2 = Ax_1 = A(Ax_0) = A^2 x_0$ etc.

So in general $x_k = A^k x_0$ for $k \geq 1$

$$A^k x_0 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n$$

$$= \lambda_1^k \left(c_1 v_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k v_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k v_n \right)$$

Since λ_1 is the dominant eigenvalue, each of these fractions go to zero as $k \to \infty$

Therefore $x_k = A^k x_0 \to \lambda_1^k c_1 v_1$

Example:

Approximate the dominant eigenvalue of $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$

Take $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as the initial vector.

$$x_1 = Ax_0 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Then $x_2 = Ax_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

k	0	1	2	3	4	5	6	7	8
x_k	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 6 \end{bmatrix}$	$\begin{bmatrix} 11 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 21 \\ 22 \end{bmatrix}$	$\begin{bmatrix} 43 \\ 42 \end{bmatrix}$	$\begin{bmatrix} 85 \\ 86 \end{bmatrix}$	$\begin{bmatrix} 171 \\ 170 \end{bmatrix}$

The iterates are clearly converging on $c\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, the dominant eigenvector.

$$Ax = \lambda x \qquad \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \lambda\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

If , then $\lambda = 2$

Practice:

Use the power method to approximate the dominant eigenvalue of $A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$ with an initial vector $x_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

THE SHIFTED POWER METHOD

Since the power method finds the dominant eigenvalue, the shifted power method understands that if λ is an eigenvalue of A, then $\lambda - \alpha A$ is an eigenvalue of $A - \alpha I$.

Continue with the previous example where $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ and we discovered that $\lambda = 2$

Then $A - 2I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$. Let $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

k	0	1	2	3	4
x_k	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1.5 \\ -3 \end{bmatrix}$	$\begin{bmatrix} 1.5 \\ -3 \end{bmatrix}$	$\begin{bmatrix} 1.5 \\ -3 \end{bmatrix}$

After 2 iterations, we find $v_2 = c\begin{bmatrix} 1 \\ -2 \end{bmatrix}$

$$Ax = \lambda x$$
$$\begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \lambda\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$
$$\begin{bmatrix} -3 \\ 6 \end{bmatrix} = \lambda\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$
$$\lambda = -3$$

INVERSE POWER METHOD

Recall that if A is invertible with eigenvalue λ, then A^{-1} has an eigenvalue $\frac{1}{\lambda}$. So if we apply the power method to A^{-1}, its dominant eigenvalue will be the reciprocal of the smallest eigenvalue of A.

Example:

We will use the inverse power method to find the second eigenvalue of $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$.

Compute $x_k = A^{-1} y_{k-1}$. Rather than computing an inverse (which takes a lot of time), we can solve the equivalent equation $Ax_k = y_{k-1}$.

Let $x_0 = y_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Solve $Ax_1 = y_0$

$$[A \mid y_0] = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Therefore $x_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, so $y_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and we get x_2 from $Ax_2 = y_1$.

$x_2 = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$, and by scaling we can get $y_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

k	0	1	2	3	4	5	6	7	8
x_k	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 1.5 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.83 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.1 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.95 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.02 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.99 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.01 \end{bmatrix}$
y_k	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -0.33 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.6 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.45 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.52 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.49 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.51 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.50 \\ 1 \end{bmatrix}$

$$Ay = \lambda y$$
$$\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$
$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$
$$\lambda = -1$$

THE SHIFTED INVERSE POWER METHOD

This can be used to find an approximation of any eigenvalue

If λ is an eigenvalue of A and $\alpha \neq \lambda$, then $A - \alpha I$ is invertible if α is not an eigenvalue of A and if $\frac{1}{(\lambda - \alpha)}$ is an eigenvalue of $(A - \alpha I)^{-1}$.

If α is close to λ, then $\frac{1}{(\lambda - \alpha)}$ will be a dominant eigenvalue of $(A - \alpha I)^{-1}$. If α is very close to λ then $\frac{1}{(\lambda - \alpha)}$ will be much bigger in magnitude than next eigenvalue, so the convergence will be very rapid.

Example: Use the shifted Power Method to find the eigenvalue closest to 5 when

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

Shifting, we would have $A - 5I = \begin{bmatrix} -5 & 5 & -6 \\ -4 & 7 & -12 \\ -2 & -2 & 5 \end{bmatrix}$.

Applying the inverse power method, let $x_0 = y_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

$$\left[A - 5I \mid y_0\right] = \left[\begin{array}{ccc|c} -5 & 5 & -6 & 1 \\ -4 & 7 & 12 & 1 \\ -2 & -2 & 5 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -0.61 \\ 0 & 1 & 0 & -0.88 \\ 0 & 0 & 1 & -0.39 \end{array}\right]$$

k	0	1	2	3	4	5	6	7
x_k	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.61 \\ -0.88 \\ -0.39 \end{bmatrix}$	$\begin{bmatrix} -0.41 \\ -0.69 \\ -0.35 \end{bmatrix}$	$\begin{bmatrix} -0.47 \\ -0.89 \\ -0.44 \end{bmatrix}$	$\begin{bmatrix} -0.49 \\ -0.95 \\ -0.48 \end{bmatrix}$	$\begin{bmatrix} -0.50 \\ -0.98 \\ -0.49 \end{bmatrix}$	$\begin{bmatrix} -0.50 \\ -0.99 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} -0.50 \\ -1.00 \\ -0.50 \end{bmatrix}$
y_k	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.69 \\ 1.00 \\ 0.45 \end{bmatrix}$	$\begin{bmatrix} 0.59 \\ 1.00 \\ 0.51 \end{bmatrix}$	$\begin{bmatrix} 0.53 \\ 1.00 \\ 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.51 \\ 1.00 \\ 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1.00 \\ 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1.00 \\ 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1.00 \\ 0.50 \end{bmatrix}$

$$Ax = \lambda x$$

$$\begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 4 \\ 8 \\ 4 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\lambda = 4$$