Matrix Methods

Matrix methods are very popular in solving problems of interests. In this lecture we are concerned with some basic numerical techniques to solve matrices.

10 Matrix Method

A large set of linear algebraic equations looks like this:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N &= b_2 \\
&\cdots \\
a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N &= b_M
\end{aligned}$$

Here N unknown x_j are related by M equations. The coefficients a_{mn}, and b_m are known numbers. In a compact form, the set of equation can be expressed in a matrix form

$$A \cdot x = b$$

where

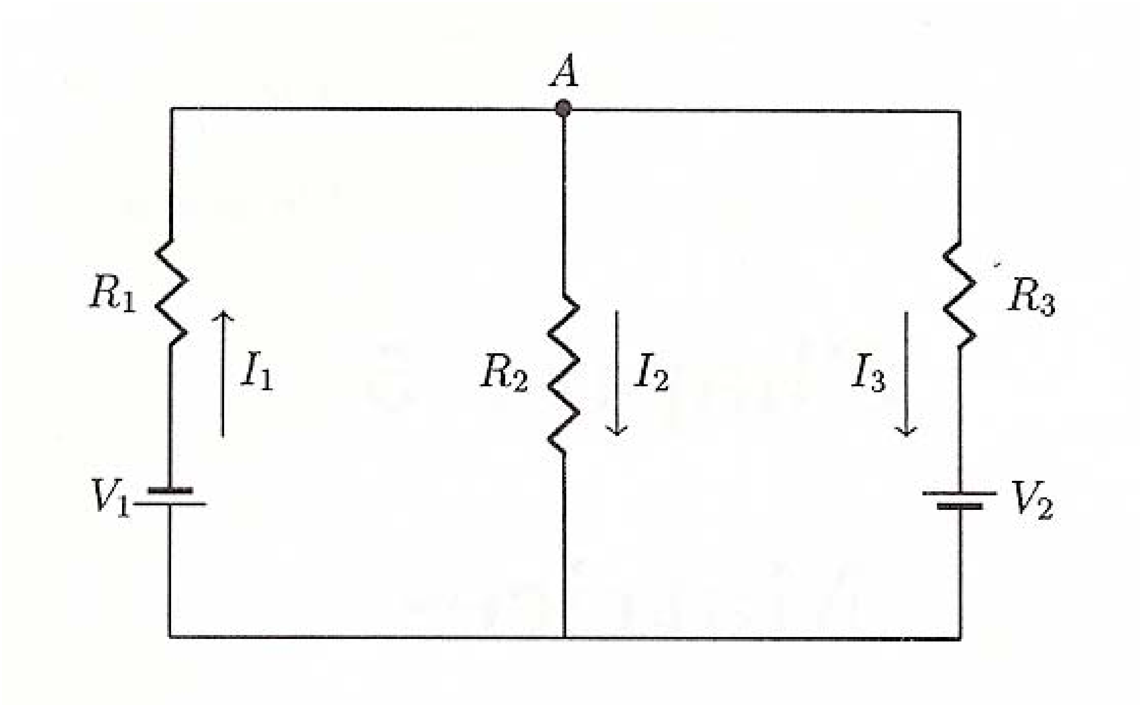
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix}.$$

By convention, the first index on an element a_{ij} denotes its row, the second index its column. If $N = M$ then there are as many equations as unknowns, and there is a good chance of solving for a unique solution set of x_i's. Analytically, there can fail to be a unique solution if one or more of M equations is a linear combination of the others, a condition called row degeneracy, or if all equations contains certain variables only in exactly the same linear combination, called column degeneracy. A set of equations that is degenerate is called singular.

10.1 Examples

10.1.1 Example 1: Direct-current circuit: the unbalanced Wheatstone bridge

We can apply the Kirchhoff rules to obtain a set of equations for the voltages and currents, and then we can solve the equation set to find the unknowns.



In a direct current circuit, we often obain a set of linear equations following the Kirchhoff rule, such as

$$R_1 I_1 + R_2 I_2 = V_1;$$

$$\begin{aligned} -R_2 I_2 + R_3 I_3 &= V_2; \\ I_1 - I_2 - I_3 &= 0. \end{aligned}$$

A more compact notation is to write the system of equations in terms of matrices,

$$\mathbf{RI} = \mathbf{V}$$

with

$$R = \begin{pmatrix} R_1 & R_2 & 0 \\ 0 & -R_2 & R_3 \\ 1 & -1 & -1 \end{pmatrix}$$

$$I = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}; V = \begin{pmatrix} V_1 \\ V_2 \\ 0 \end{pmatrix}$$

10.1.2 Example 2: The vibrational spectrum of molecula

For a simple harmonic oscillator,

$$m\ddot{x} = -kx$$

Consider a molecular system with multiple degrees of freedom.

Te potential energy:

$$U = \frac{1}{2}\sum_{j,k}^{n} A_{jk} q_j q_k$$

The kinetic energy:

$$T = \frac{1}{2}\sum_{j,k}^{n} M_{jk} \dot{q}_j \dot{q}_k$$

The Lagrangian equation:

$$\frac{\partial L}{\partial q_j} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{q}_j} = 0$$

with $L = T - U$ the Lagrangian of the system,

$$\sum_j^n \left(A_{jk} q_j + M_{jk} \ddot{q}_j \right) = 0.$$

If we assume the time dependence of the generalized coordinates is oscillatory with an angular frequency ω, $q_j = x_j \exp(i\omega t)$, we have

$$\sum_j^n \left(A_{jk} - M_{jk}\omega^2 \right) x_j = 0.$$

In the form of matrix

$$\left(\mathbf{A} - \mathbf{M}\omega^2 \right) \cdot x = 0$$

The vibrational frequencies ω_k with $k = 1, 2, \cdots, n$ are then obtained by solving the above secular equation.

Example: the vibrational spectrum of one-dimensional crystal (with an impurity)

10.1.3 Example 3: The Heisenberg model:

$$H = J \sum_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$$

For a two-spin 1/2 system, it can be reduced to a 4 × 4 matrix issue.

$$\begin{aligned}
\phi_1 &= |1,1\rangle \equiv |1\rangle \otimes |1\rangle \, , \\
\phi_2 &= |1,-1\rangle \, , \\
\phi_3 &= |-1,1\rangle \, , \\
\phi_4 &= |-1,-1\rangle \, ,
\end{aligned}$$

For a three-spin 1/2 system, it can be reduced to a 8×8 matrix issue. If you consider the symmetry in the Hamiltonian, you may simplify the problem.

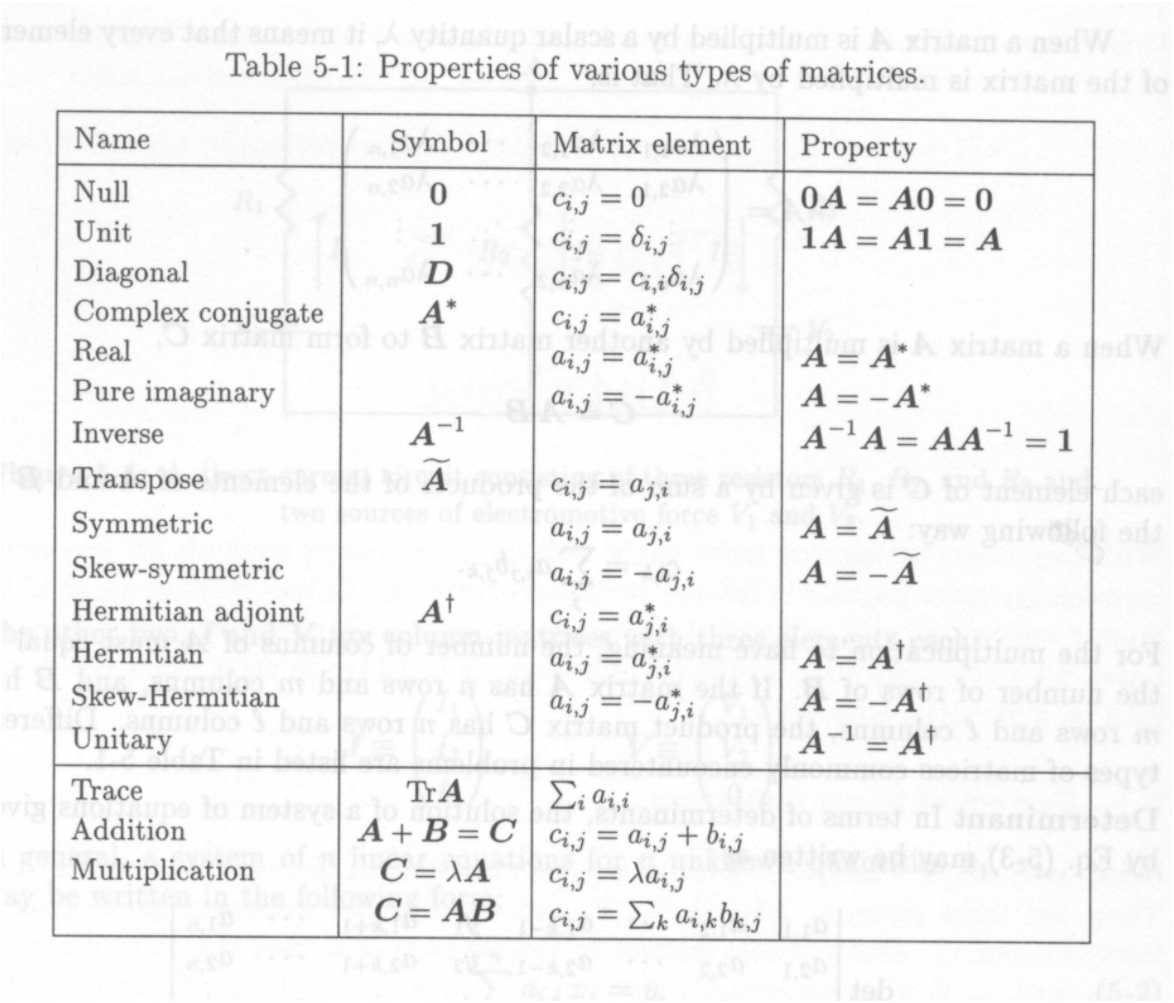


Table 5-1: Properties of various types of matrices.

Name	Symbol	Matrix element	Property
Null	$\mathbf{0}$	$c_{i,j} = 0$	$\mathbf{0A} = \mathbf{A0} = \mathbf{0}$
Unit	$\mathbf{1}$	$c_{i,j} = \delta_{i,j}$	$\mathbf{1A} = \mathbf{A1} = \mathbf{A}$
Diagonal	$\mathbf{D}$	$c_{i,j} = c_{i,i}\delta_{i,j}$	
Complex conjugate	$\mathbf{A}^*$	$c_{i,j} = a^*_{i,j}$	
Real		$a_{i,j} = a^*_{i,j}$	$\mathbf{A} = \mathbf{A}^*$
Pure imaginary		$a_{i,j} = -a^*_{i,j}$	$\mathbf{A} = -\mathbf{A}^*$
Inverse	$\mathbf{A}^{-1}$		$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{1}$
Transpose	$\widetilde{\mathbf{A}}$	$c_{i,j} = a_{j,i}$	
Symmetric		$a_{i,j} = a_{j,i}$	$\mathbf{A} = \widetilde{\mathbf{A}}$
Skew-symmetric		$a_{i,j} = -a_{j,i}$	$\mathbf{A} = -\widetilde{\mathbf{A}}$
Hermitian adjoint	$\mathbf{A}^\dagger$	$c_{i,j} = a^*_{j,i}$	
Hermitian		$a_{i,j} = a^*_{j,i}$	$\mathbf{A} = \mathbf{A}^\dagger$
Skew-Hermitian		$a_{i,j} = -a^*_{j,i}$	$\mathbf{A} = -\mathbf{A}^\dagger$
Unitary			$\mathbf{A}^{-1} = \mathbf{A}^\dagger$
Trace	$\mathrm{Tr}\mathbf{A}$	$\sum_i a_{i,i}$	
Addition	$\mathbf{A} + \mathbf{B} = \mathbf{C}$	$c_{i,j} = a_{i,j} + b_{i,j}$	
Multiplication	$\mathbf{C} = \lambda\mathbf{A}$	$c_{i,j} = \lambda a_{i,j}$	
	$\mathbf{C} = \mathbf{AB}$	$c_{i,j} = \sum_k a_{i,k}b_{k,j}$	

Figure 1:

10.2 Definition and basic facts

10.3 Tasks of computational linear algebra

- Solution of matrix equation $Ax = b$ for an unknown vector x, where A is a square matrix of coefficients.

- Solution of more than one matrix equation $Ax_j = b_j$ for a set of vectors x_j, $j = 1, 2, \cdots$, each corresponding to a different, known right-hand side vector b_j.

- Calculation of the matrix A^{-1} which is the matrix inversion of a square matrix A.

- Calculation of the determinant of a square matrix A.

- Eigenvalues and eigenvectors of a square matrix A.

11 System of linear equations

11.1 Determinant

In terms of determinants, the solution of a linear system may be written by

$$x_k = \frac{\det \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k-1} & b_1 & a_{1k+1} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2k-1} & b_2 & a_{2k+1} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{Nk-1} & b_N & a_{Nk+1} & \cdots & a_{NN} \end{vmatrix}}{\det \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix}}$$

where the denominator is the determinant of A itself and the numerator is that of A with the element in column k replaced by those of b. In this case our problem is reduced to one of evaluating determinants. For example, for $N = 2$,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21};$$

for $N = 3$,

$$\begin{aligned} & \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ = \; & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} \\ & -a_{13}a_{22}a_{31} - a_{32}a_{23}a_{11} - a_{12}a_{21}a_{33} \end{aligned}$$

The value of the determinant of a square matrix is given by

$$\det A = \sum \epsilon_{i,j,k,\cdots} a_{1,i} a_{2,j} a_{3,k} \cdots$$

Here $\epsilon_{i,j,k,\cdots}$ are the Levi-Civita symbols. It is equal to $+1$ for any even permutation of the n subscripts $\{$1,2,3,4,$\cdots$,n$\}$, -1 for an odd permutation, and zero if any two indices are the same.

In general, one way to evaluate a determinant is based on the Laplace expansion theorem

$$\det A = \sum_i (-1)^{i+j} a_{ij} M_{ij}$$

where the minor M_{ij} is defined as a determinant obtained from A by removing row i and column j. However, this method becomes very cumbersome if N is very large. For example,

$$M_{22} = \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}$$

11.2 Gauss-Jordan elimination method

Property of a determinant: its value of a determinant is unchanged if a column (or row) is replaced by a linear combination of itself and other columns (or row). For example,

$$\det \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix} = \det \begin{vmatrix} a_{11} + \lambda a_{12} & a_{12} & \cdots & a_{1N} \\ a_{21} + \lambda a_{22} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} + a_{N2} & a_{N2} & \cdots & a_{NN} \end{vmatrix}$$

We can make use of this property to transform a determinant such that all off-diagonal elements vanish. For example, assume $a_{11} \neq 0$. By subtracting

a_{21}

$$
\begin{aligned}
& \det\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix} \\
= \; & \det\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} - \frac{a_{11}a_{21}}{a_{11}} & a_{22} - \frac{a_{12}a_{21}}{a_{11}} & \cdots & a_{2N} - \frac{a_{1N}a_{21}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} - \frac{a_{11}a_{N1}}{a_{11}} & a_{N2} - \frac{a_{12}a_{N1}}{a_{11}} & \cdots & a_{NN} - \frac{a_{1N}a_{N1}}{a_{11}} \end{vmatrix} \\
= \; & \det\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & a'_{22} & \cdots & a'_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{N2} & \cdots & a'_{NN} \end{vmatrix}
\end{aligned}
$$

i.e.,

$$a'_{ij} = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}$$

for $j = 1, 2, \cdots, N$ and $i = 2, 3, \cdots, N$.

Then perform the transform further,

$$a'_{ij} = a_{ij} - \frac{a_{i2}a_{2j}}{a_{22}}$$

for $j = 2, \cdots, N$ and $i = 3, \cdots, N$. We have

$$\det\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix} = \det\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & a'_{22} & \cdots & a'_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a'_{NN} \end{vmatrix}$$

Repeat the transform

$$a'_{ij} = a'_{ij} - \frac{a'_{ik}a'_{kj}}{a'_{22}}$$

for $j = k, \cdots, N$ and $i = k+1, \cdots, N$ until $k = N-1$. The value of determinant is equivalent to that of an upper triangle matrix and its value is determined by the product of diagonal elements

$$\det A = a_{11}a'_{22}a'_{33}\cdots a'_{NN}.$$

For example

$$\begin{aligned}
\det\begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{vmatrix} &= \det\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1-2\times 3 & 2-3\times 3 \\ 0 & 3-2\times 2 & 1-3\times 2 \end{vmatrix} \\
= \det\begin{vmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{vmatrix} &= \det\begin{vmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -5+7\times\frac{1}{5} \end{vmatrix} \\
= 1\times(-5)\times(-\frac{18}{5}) &= 18
\end{aligned}$$

The algebra of this method to evaluate the value of determinant is very simple. In practice, a slight refinement is needed. The basic operation in Gaussian elimination is to reduce the off-diagonal element at position (i, j)

to zero by subtracting from its multiples of a diagonal element. Such a diagonal element is usually referred to as the pivot. This methods fails if, for any reason, one of the pivot vanishes. By the same token, the numerical accuracy of the result will be poor if some of the pivots used are much smaller compared with the value of the off-diagonal elements involved in the same calculation. To prevent this from happening, it is advantageous to use as pivot the element with largest possible absolute value. In practice, this principle is applied in the following way. Consider the situation that the determinant is transferred to the stage that all the off-diagonal matrix elements in the lower half-triangle up to row k are reduced to zero. A search is made among the $(N-k)\times(N-k)$ remaining element to locate the one with the largest absolute value. If this element is not at the position of next pivot, the diagonal element at position (k,k), we shall permute the rows and columns of the determinant so as to bring it to position (k,k). This method is known as Gauss-Jordan elimination.

11.3 Solution of linear equations by elimination

Instead of evaluating the determinants, it is also possible to solve a system of linear equations directly by Gauss-Jordan elimination. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 11 \\ 11 \end{pmatrix}$$

We subtract three times of the first row from the second row to eliminate the first element in the second row, and two time of the first row from the third row.

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ -31 \\ -17 \end{pmatrix}$$

We subtract 1/5 time of the second row from the third row

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -\frac{18}{5} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ -31 \\ -\frac{54}{5} \end{pmatrix}$$

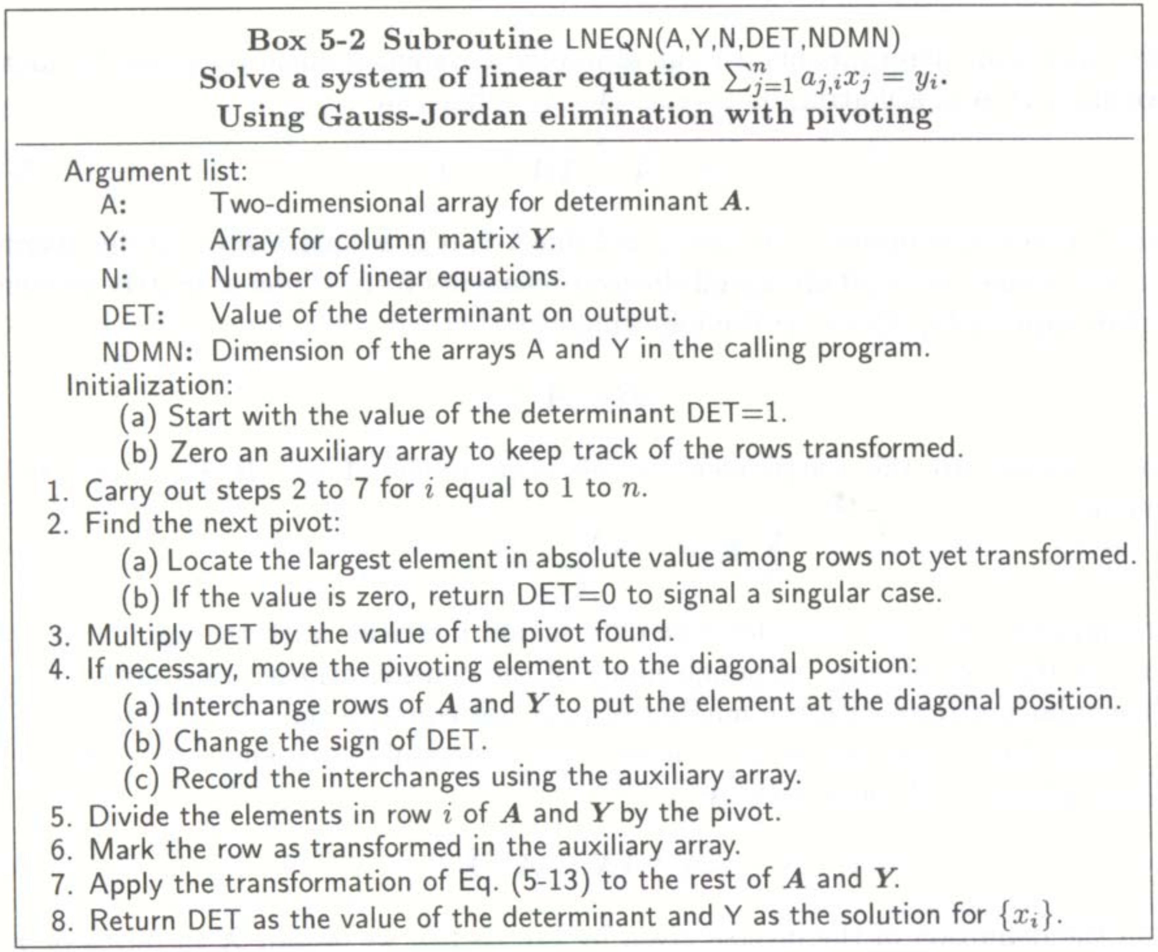
From the third row, we have the solution, $x_3 = 3$. Inserting x_3 into the second row, we obtain $x_2 = 2$. Then, $x_1 = 1$.

To check the result, we can substitute the solution obtained for $\{x_i\}$ into the original linear equations and see if the sum of the product $\sum_j a_{i,j} x_j = y_i$ within the accuracies required.

Problem: For a list of data, $\{a_1, a_2, \cdots, a_n\}$,

1. write a program to find the maximal value of a_i;

2. re-organize the list with an descent order.



Box 5-2 Subroutine LNEQN(A,Y,N,DET,NDMN)
Solve a system of linear equation $\sum_{j=1}^{n} a_{j,i} x_j = y_i$.
Using Gauss-Jordan elimination with pivoting

Argument list:

A:	Two-dimensional array for determinant $\boldsymbol{A}$.
Y:	Array for column matrix $\boldsymbol{Y}$.
N:	Number of linear equations.
DET:	Value of the determinant on output.
NDMN:	Dimension of the arrays A and Y in the calling program.

Initialization:

(a) Start with the value of the determinant DET=1.
(b) Zero an auxiliary array to keep track of the rows transformed.

1. Carry out steps 2 to 7 for i equal to 1 to n.
2. Find the next pivot:
   (a) Locate the largest element in absolute value among rows not yet transformed.
   (b) If the value is zero, return DET=0 to signal a singular case.
3. Multiply DET by the value of the pivot found.
4. If necessary, move the pivoting element to the diagonal position:
   (a) Interchange rows of $\boldsymbol{A}$ and $\boldsymbol{Y}$ to put the element at the diagonal position.
   (b) Change the sign of DET.
   (c) Record the interchanges using the auxiliary array.
5. Divide the elements in row i of $\boldsymbol{A}$ and $\boldsymbol{Y}$ by the pivot.
6. Mark the row as transformed in the auxiliary array.
7. Apply the transformation of Eq. (5-13) to the rest of $\boldsymbol{A}$ and $\boldsymbol{Y}$.
8. Return DET as the value of the determinant and Y as the solution for $\{x_i\}$.

Figure 2:

12 Matrix inversion

Instead of using determinants, we can solve a system of linear equations by matrix inversion. If A is a matrix, its inversion is defined as

$$A \cdot A^{-1} = A^{-1} \cdot A = 1.$$

Let $B = A^{-1}$. Then

$$\sum_k b_{i,k} a_{k,j} = \delta_{ij}$$

The solution of a system of linear equations can be also written by

$$x = A^{-1} b.$$

A^{-1} is the inverse of the matrix A:

$$A \cdot A^{-1} = A^{-1} \cdot A = 1.$$

Let $B = A^{-1}$. We have

$$x_i = \sum_j b_{ij} b_j.$$

In principle, we can find the inverse matrix by making use of its definition. In this way we have to solve a total of n^2 equations to find the n^2 unknown elements b_{ij}.

We shall regard this merely as a statement that the problem has a solution. The actual calculations involved are much simpler. In practice, the solution requires only a single application of a slightly modified form of Gaussian-Jordan elimilation, as we shall see next.

12.1 Gauss-Jordan elimination

To find the inverse of matrix A, we can use the Gauss-Jordan elimination. We have applied the Gauss-Jordan method to evaluate the determinant. For our discussion here, the operation is expressed symbolically in terms of an operator

$$\mathcal{O}A = \mathsf{det}\,|A|\,1$$

where $\mathcal{O}$ represents all the steps required to reduce A to a unit matrix. The same operation on a unit matrix gives

$$\mathcal{O}A\left(\leftarrow A^{-1}\right) = \mathsf{det}\,|A|\,1\left(\leftarrow A^{-1}\right)$$

$$\mathcal{O}1 = A^{-1}\,\mathsf{det}\,|A|\,.$$

The basic idea can be seen from the following example

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & 1 & 2 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -3 \\ 3 & -5 & -7 & 0 & 1 & 0 \\ 2 & -1 & -5 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{5} & \frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & 0 & \frac{3}{5} & -\frac{1}{5} & -\frac{7}{5} \\ \frac{7}{5} & \frac{1}{5} & -\frac{18}{5} & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{5}{18} & \frac{7}{18} & \frac{1}{18} \\ 0 & 1 & 0 & \frac{1}{18} & -\frac{5}{18} & \frac{7}{18} \\ 0 & 0 & 1 & \frac{7}{18} & \frac{1}{18} & -\frac{5}{18} \end{array}\right)$$

From a practical point of view, we can interpret Gauss-Jordan elimilation in this way. The inverse of a matrix A may be found by the same steps as used above to reduce it to a unit matrix. To store the steps, we can apply the operations to a unit matrix at the same time. Except for a normalization constant, given by the value of determinant, A is transformed at the end into a unit matrix, while the unit matrix is transformed into the inverse of A.

Two improvements: 1). reducing the storage required for the elements; 2). Pivoting.

12.2 LU-decomposition

12.2.1 Triangular Systems: forward substitution

$$L = \begin{pmatrix} l_{1,1} & 0 & 0 & \cdots & 0 \\ l_{2,1} & l_{2,2} & 0 & \cdots & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ l_{n,1} & l_{n,2} & l_{n,3} & \cdots & l_{n,n} \end{pmatrix}$$

Consider the following 2 by 2 lower triangular systems:

$$\begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

If $l_{11}l_{22} \neq 0$, then the unknowns are determined by

$$\begin{aligned} x_1 &= b_1/l_{11} \\ x_2 &= (b_2 - l_{21}x_1)/l_{22} \end{aligned}$$

The general procedure is obtained by solving the ith equation in $Lx = b$ for x_i,

$$x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{j=1}^{i-1} l_{ij} x_j \right).$$

12.2.2 Triangular Systems: Backward substitution

$$L = \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \cdots & u_{1,n} \\ 0 & u_{2,2} & u_{2,3} & \cdots & u_{2,n} \\ 0 & 0 & u_{3,3} & \cdots & u_{3,n} \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & u_{n,n} \end{pmatrix}$$

Consider a 2 × 2 upper triangle systems

$$\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

If $u_{11}u_{22} \neq 0$, then the unknowns are determined by

$$\begin{aligned} x_2 &= b_2/u_{22} \\ x_1 &= (b_1 - u_{12}x_2)/u_{11} \end{aligned}$$

For upper triangle systems $Ux = b$, the solution can be written by

$$x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{j=i+1}^{N} u_{ij} x_j \right) .$$

12.2.3 LU-decomposition

Since the triangular systems are “easy” to solve, we shall try to take advantage of these properties in solving a system of linear equations $Ax = b$. Assume the matrix A can be written as a product of an upper triangular matrix U and a lower triangular matrix L: $A = LU$ where

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{N1} & l_{N2} & \cdots & l_{NN} \end{pmatrix}$$

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ 0 & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{NN} \end{pmatrix}$$

In terms of L and U, the equation may be written as

$$Ax = L(Ux) = Lz = b$$

We write

$$\begin{aligned} Lz &= b \\ Ux &= z \end{aligned}$$

We can use the forward substitution to obtain z_i, and then use the back substitution to obtain x_i.

We are now in the position how to decompose the matrix as a product of a lower triangular matrix and an upper triangular matrix: $A = LU$.

From $A = LU$, we have N^2 equations. L and U each has $N(N+1)/2$ unknowns, and totally $N^2 + N$. So we have the freedom to assign values to N elements of L (or U). We choose

$$l_{ii} = 1, \text{ for } i = 1, \cdots, N.$$

The elements of L, U and A are related by

$$\sum_m l_{im} u_{mj} = a_{ij}$$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{N1} & l_{N2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ 0 & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{NN} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix}$$

Since the particular form of triangular matrices,

$$l_{ij} = 0,$$

for all $i < j$;

$$u_{ij} = 0$$

for all $i > j$ we can work out the unknown l_{ij} and u_{ij} in a simple way. For $i = 1$,

$$u_{1j} = a_{1j},$$

for all j. This gives the first row of U.

For $j = 1$, we have

$$l_{i1}u_{11} = a_{i1} \rightarrow l_{i1} = a_{i1}/u_{11}$$

for $i = 2, \cdots, N$. This gives the firs column of L.

With all the elements in the first column known, we can proceed to solve for those in the second column.

$$l_{2,1}u_{1,j} + u_{2,j} = a_{2,j}$$

$$u_{2,j} = a_{2,j} - l_{2,1}u_{1,j}$$

Similarly,

$$l_{3,1}u_{1,j} + l_{3,2}u_{2,j} + l_{3,3}u_{3,j} = a_{3,j}$$

$$l_{3,3}u_{3,j} = a_{3,j} - (l_{3,1}u_{1,j} + l_{3,2}u_{2,j})$$

The general expressions for l_{ij} and u_{ij} are, respectively,

$$u_{ij} = a_{ij} - \sum_{m=1}^{i-1} l_{im}u_{mj}$$

for $i = 1, 2, \cdots, k$.

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{m=1}^{j-1} l_{im} u_{mj} \right)$$

for $i = 1, 2, \cdots, k-1$.

Probem in class: Decompose a 3×3 matrix in LU method.

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

Pivoting: It is important to note here that each step of the calculation requires the result of the previous step. For example, the element $u_{\alpha,k}$ is found from the values of $u_{m,k}$ for $m < \alpha$, as well as the elements of L

in the earlier columns. Similarly, to calculate $l_{\alpha,k}$, we need the values of the elements of U in the same column, as well as the those of L in the previous columns. In calculations of this type, the stability of solution can be a problem and the errors in each step become cumulative in such a way that hardly any significant figures are left at the end. Pivoting is one way to improve the accuracy and stabilize the solution.

Example:

$$A = \begin{pmatrix} 0.0001 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 10,000 & 1 \end{pmatrix} \begin{pmatrix} 0.0001 & 1 \\ 0 & -9999 \end{pmatrix}$$

It correctly identifies the source of the difficulty: relatively small pivots. A way out of this difficulty is to interchange rows. In this example, if the P is the permutation

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

then

$$PA = \begin{pmatrix} 1 & 1 \\ 0.0001 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.0001 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0.9999 \end{pmatrix}$$

Now the triangle factors are comprised of acceptably small elements.

Storage information: Once all elements of a particular column of U and L are calculated, the elements in the same column of A are no longer needed. Since the calculation is carried out in the order from left to right one column at a time, and for each column, from the top to the bottom one element at a time, the elements of A can be discarded one after another as the reduction of A to L and U progresses. The array for A may therefore be used to store both U and L.

13 The eigenvalue problem

13.1 The eigenvalue problem

In quantum mechanics, the starting point of studying many problems is the time independent Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi_\alpha = E_\alpha\psi_\alpha.$$

There are several ways to solve this equation. One of them is by matrix method.

Basis states

The first step in taking a matrix approach is to choose a complete and orthogonal set of states $\{\phi_1, \phi_2, \phi_3, \cdots \phi_n\}$ such that all states can be expanded in terms of these basis states.

$$\psi_\alpha = \sum_{i=1}^{n} c_{\alpha i}\phi_i$$

The basis states is normalized and orthogonal,

$$\int \phi_i^* \phi_j dr = \delta_{ij}$$

Multiplying the complex conjugate of a basis state on the both sides of equation, we have

$$\int \phi_i^* \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi_\alpha dr = \int E_\alpha \phi_i^* \psi_\alpha dr$$

$$\sum H_{ij} c_{\alpha j} = E_\alpha c_{\alpha i}$$

where

$$H_{ij} = \int \phi_i^* \left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) \phi_j dr$$

In a matrix form,

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{nn} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} \end{pmatrix} \begin{pmatrix} c_{\alpha 1} \\ c_{\alpha 2} \\ \vdots \\ c_{\alpha n} \end{pmatrix} = E_\alpha \begin{pmatrix} c_{\alpha 1} \\ c_{\alpha 2} \\ \vdots \\ c_{\alpha n} \end{pmatrix}$$

The eigenvalues are determined by the characteristic equation

$$\det \begin{vmatrix} H_{11} - E_\alpha & H_{12} & \cdots & H_{nn} \\ H_{21} & H_{22} - E_\alpha & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} - E_\alpha \end{vmatrix} = 0$$

If we choose the energy eigenstates as the basis state,

$$\tilde{H}_{\alpha\beta} = E_\alpha \delta_{\alpha\beta}$$

where

$$\tilde{H}_{\alpha\beta} = \int \psi_\alpha^* \left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) \psi_\beta dr$$

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix}$$

We have

$$\left(U^\dagger H U \right)_{\alpha\beta} = E_\alpha \delta_{\alpha\beta}$$

Since $U^\dagger = U^{-1}$, this is a similarity transform, i.e., the Hamiltonian can be reduced to a diagonal matrix under a similarity transformation.

13.2 Two-dimensional rotation

The most straightforward method to diagonalize a real symmetric matrix is the Jacobi method. In this approach, the basis operation consists of a series of rotations among two column and two rows. Consider a real symmetric 2×2 matrix of the form

$$A = \left(\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right)$$

with $a_{12} = a_{21}$. We write a transformation matrix in terms of trigonometric function in the following form

$$T = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

The matrix is unitary,

$$T^\dagger = T^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Under the transformation,

$$T^{-1}AT = \begin{pmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{pmatrix}$$

where

$$\begin{aligned} a'_{11} &= a_{11}\cos^2\theta + a_{22}\sin^2\theta - a_{12}\cos\theta\sin\theta \\ a'_{22} &= a_{11}\sin^2\theta + a_{22}\cos^2\theta + a_{12}\cos\theta\sin\theta \\ a'_{12} &= (a_{11} - a_{22})\cos\theta\sin\theta + (\cos^2\theta - \sin^2\theta)a_{12} \end{aligned}$$

To find a diagonal matrix, the off-diagonal elements should be equal to zero,

$$\frac{2\cos\theta\sin\theta}{\cos^2\theta - \sin^2\theta} = \frac{-2a_{12}}{a_{11}-a_{22}} = \tan 2\theta$$

Its solution is

$$\begin{aligned}
\sin 2\theta &= \frac{2a_{12}}{\left[4a_{12}^2 + (a_{11}-a_{22})^2\right]^{1/2}} \\
\cos 2\theta &= \frac{a_{22}-a_{11}}{\left[4a_{12}^2 + (a_{11}-a_{22})^2\right]^{1/2}}
\end{aligned}$$

The diagonal elements are

$$\begin{aligned}
a'_{11} &= \frac{a_{11}+a_{22}}{2} - \frac{1}{2}\left[4a_{12}^2 + (a_{11}-a_{22})^2\right]^{1/2} \\
a'_{22} &= \frac{a_{11}+a_{22}}{2} + \frac{1}{2}\left[4a_{12}^2 + (a_{11}-a_{22})^2\right]^{1/2}
\end{aligned}$$

Eigenvectors:

$$\begin{aligned} v_1 &= \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix} \\ v_2 &= \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} \end{aligned}$$

To evaluate $\cos\theta$ and $\sin\theta$, from $\cos 2\theta$

$$\begin{aligned} \cos\theta &= \pm\left(\frac{1}{2} + \frac{1}{2}\cos 2\theta\right)^{1/2} \\ \sin^2\theta &= \pm\left(\frac{1}{2} - \frac{1}{2}\cos 2\theta\right)^{1/2} \end{aligned}$$

and from $\sin 2\theta = 2\cos\theta\sin\theta$,

$$\cos\theta = \frac{1}{2^{1/2}}\left(1 + \frac{a_{22} - a_{11}}{\left(4a_{12}^2 + (a_{22} - a_{11})^2\right)^{1/2}}\right)^{1/2}$$

$$\sin\theta \;=\; \frac{sgn(a_{12})}{2^{1/2}}\left(1-\frac{a_{22}-a_{11}}{\left(4a_{12}^2+(a_{22}-a_{11})^2\right)^{1/2}}\right)^{1/2}$$

In general, if

$$V^{-1}AV = D = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

then the column vector

$$v_m = \begin{pmatrix} v_{1m} \\ v_{2m} \\ \vdots \\ v_{nm} \end{pmatrix}$$

is the eigen vector of A if

$$Av_m = \lambda_m v_m.$$

13.3 Jacobi method

The same type of rotation may be applied to large matrices. Consider a real symmetric matrix A of some finite dimension n. If we wish to put to zero a particular pair of off-diagonal matrix elements, for example, a_{kl} and a_{lk} we can apply a two dimensional rotation. The major difference is that, instead of two dimensional, the transformation matrix is n-dimensional. To bring the complete matrix to a diagonal form, the Jacobi method makes a series of such rotation, each of which annihilates a pair of off-diagonal matrix elements. Let us assume that after $m-1$ such a rotation, the original matrix is transformed into $A(m-1)$ with matrix elements $\left\{a_{i,j}^{(m-1)}\right\}$. Since all the transformation is unitary, the matrix $A(m-1)$ remains real and symmetric. Let us further assume that, in the next rotation, we wish to annihilate the pair of elements $a_{k,l}^{(m-1)}$ and $a_{l,k}^{(m-1)}$, the transformation matrix is written as

$$T = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \cos\theta^{(m)} & \cdots & \sin\theta^{(m)} & 0 \\ \vdots & \vdots & 1 & \vdots & \vdots \\ 0 & -\sin\theta^{(m)} & \cdots & \cos\theta^{(m)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

That is, (1) all diagonal elements are 1 except

$$t_{kk} = t_{ll} = \cos\theta^{(m)}$$

(2) all off-diagonal matrix elements are zero except

$$t_{kl} = -t_{lk} = \sin\theta^{(m)}$$

The angle $\theta^{(m)}$ is determined by

$$\begin{aligned}\sin\theta^{(m)} &= \frac{a_{12}^{(m-1)}}{\left[a_{12}^{(m-1)}a_{12}^{(m-1)} + \left(a_{11}^{(m-1)} - a_{22}^{(m-1)}\right)^2\right]^{1/2}}\\ \cos\theta^{(m)} &= \frac{a_{22}^{(m-1)} - a_{11}^{(m-1)}}{\left[a_{12}^{(m-1)}a_{12}^{(m-1)} + \left(a_{11}^{(m-1)} - a_{22}^{(m-1)}\right)^2\right]^{1/2}}\end{aligned}$$

The elements of the transformed matrix are

(1) the elements outside the column k and row l are unchanged by the transformation

$$a_{i,j}^{(m)} = a_{i,j}^{(m-1)} \text{ for i} \neq k, l,\ \ j \neq k, l$$

(2) the elements in column k, and row l

$$\begin{aligned}
a_{kk}^{(m)} &= a_{kk}^{(m-1)} \cos^2\theta + a_{ll}^{(m-1)} \sin^2\theta - a_{kl}^{(m-1)} \cos\theta \sin\theta \\
a_{ll}^{(m)} &= a_{kk}^{(m-1)} \sin^2\theta + a_{ll}^{(m-1)} \cos^2\theta + a_{kl}^{(m-1)} \cos\theta \sin\theta \\
a_{kl}^{(m)} &= a_{lk}^{(m)} = 0.
\end{aligned}$$

(3) other elements in column k and row l

$$\begin{aligned}
a_{ki}^{(m)} &= a_{ik}^{(m)} = a_{ik}^{(m-1)} \cos\theta - a_{il}^{(m-1)} \sin\theta \\
a_{li}^{(m)} &= a_{il}^{(m)} = a_{ik}^{(m-1)} \sin\theta + a_{il}^{(m-1)} \cos\theta.
\end{aligned}$$

for $i \neq k, l$.

Because the third group of transformation in the matrix elements, the Jacobi method must be iterative. This can be seen from the fact that the off-diagonal matrix elements that vanished in an earlier transformation may become nonzero by a subsequent transformation. The process can, however, be shown to be convergent, since successive transformation increase the absolute values of the diagonal elements and decrease the absolute values of off-diagonal elements. We shall not prove it here, and just give an example for a 2 × 2 matrix. Define a quantity

$$off(A) = \left(\sum_{i=1}^{n} \sum_{j \neq i} a_{i,j}^2 \right)^{1/2} .$$

We have the following inequality,

$$off(A(m)) \leq off(A(m-1)).$$

The equal sign holds for all off-diagonal elements are equal to zero. For a 2 × 2 matrix,

$$off(\tilde{A}) = off(A) - a_{12}a_{21} < off(A).$$

It is very trivial.

Eigenstates:

$$D = V^{-1}AV$$

The eigenstate of D:

$$\epsilon_i = \begin{pmatrix} \vdots \\ 1 \\ \vdots \end{pmatrix}$$

The eigenstate of A is

$$\phi_i = V\epsilon_i$$

Improving efficiency: For computational efficiency, it is advantageous to avoid calculating the trigonometry function explicitly.

14 Tri-diagonalization method of Givens and Householder

The following discussion is restricted to real symmetric matrices.

14.1 Inner product and outer product

Before we discuss the reduction of a matrix to a tri-diagonal matrix, it is useful to review some properties of matrix multiplication. A column vector with n elements

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

may be thought as a matrix consisting of a single row of n elements. Its transpose is

$$a^T = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}.$$

Inner product:

$$a^T b = \sum a_i b_i$$

The inner product of a and a^T is given by the square of norm of the vector

$$a^T a = \sum a_i^2 = ||a||^2$$

Outer product:

$$ba^T = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

$$= \begin{pmatrix} a_1b_1 & a_2b_1 & \cdots & a_nb_1 \\ a_1b_2 & a_2b_2 & \cdots & a_nb_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1b_n & a_2b_n & \cdots & a_nb_n \end{pmatrix}$$

If you learnt the Dirac notation in quantum mechanics, the ket $|a\rangle$ and bra $\langle b|$, the inner and outer product are $\langle a|b\rangle$ and $|a\rangle \langle b|$, respectively.

14.2 Method of Householder

This is a method to reduce a square matrix to a tri-diagonal one. Consider the following matrix

$$P = 1 - \beta \, \mathsf{u} \, \mathsf{u}^T$$

where

$$u = a \pm ||a||\,\varepsilon_1$$

and

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\beta = \frac{2}{||u||^2}$$

we have

$$Pa = \pm ||a||\,\varepsilon_1$$

$$
\begin{aligned}
Pa &= \left(1 - \beta\, \mathsf{u}\, \mathsf{u}^T\right) a \\
&= a - \beta\, \mathsf{u}(a^T \pm ||a||\, \varepsilon_1^T)a \\
&= a - \beta u(a^T a \pm ||a||\, \varepsilon_1^T a) \\
&= a - \beta u(||a||^2 \pm ||a||\, a_1)
\end{aligned}
$$

The norm of u is

$$
\begin{aligned}
||u||^2 &= (a^T \pm ||a||\, \varepsilon_1^T)(a \pm ||a||\, \varepsilon_1) \\
&= 2\, ||a||\, (||a|| \pm a_1)
\end{aligned}
$$

As a result, if we take

$$\beta = \frac{2}{||u||^2},$$

we got the indentity. Since all the elements of ε_1 are zero except the first row, only the first element of Pa is different from zero.

Reduction of one row and one column to tri-diagonal form

For a matrix A, we define

$$T(1) = 1 - \beta_1 w^{(1)} w^{(1)T}$$

where

$$w^{(1)} = \begin{pmatrix} 0 \\ a_{21} \pm \alpha \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}$$

$$\begin{aligned} \alpha^2 &= \sum_{i=2}^{n} a_{i1}^2 \\ \beta_1 &= \frac{2}{\left\|w^{(1)}\right\|^2} \end{aligned}$$

The $T(1)$ is unitary

$$T^{-1}(1) = T^{T}(1) = T(1)$$

Under the transformation of $T(1)$, we have

$$A(1) = T(1)AT(1) = \begin{pmatrix} a_{11} & \pm\alpha & 0 & \cdots & 0 \\ \pm\alpha & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}$$

We proceed to apply a second transformation

$$T(2) = 1 - \beta_2 w^{(2)} w^{(2)T}$$

where

$$w^{(2)} = \begin{pmatrix} 0 \\ 0 \\ a'_{32} \pm \alpha \\ \vdots \\ a'_{n2} \end{pmatrix}$$

$$A(2) = T(2)A(1)T(2)$$

In general,

$$A(m) = T(m)A(m-1)T(m)$$

$$T(m) = 1 - \beta_m w^{(m)} w^{(m)T}$$

where

$$w_i^{(m)} = \begin{cases} 0 & \text{for} \quad i \leq m \\ a_{m+1,m}^{(m-1)} \pm \alpha & \text{for} \quad i = m+1 \\ a_{m,i}^{(m-10)} & \text{for} \quad m+1 < i \leq n \end{cases}$$

After $n-1$ such transforms, we arrive at a tri-diagonal form.

$$R^T A R = J$$

where

$$R = T(1)T(2)\cdots T(n-1)$$

It is also possible to start the tridiagonalization process from the last column and row, instead of the first column and row we used.

15 Eigenvalues and eigenvectors of a tri-diagonal matrix

The most straightforward method to find the eigenvalues of a real symmetric tridiagonal matrix J is to solve the characteristic equation,

$$\det(J - \lambda 1) = 0$$

15.1 Sturm sequence of polynomials

For $n = 2$,

$$p_2(\lambda) \;=\; \det\begin{vmatrix} d_1 - \lambda & f_2 \\ f_2 & d_2 - \lambda \end{vmatrix}$$

$$= \quad (d_1 - \lambda)(d_2 - \lambda) - f_2^2$$

For $n = 3$,

$$\begin{aligned} p_3(\lambda) &= \det \begin{vmatrix} d_1 - \lambda & f_2 & 0 \\ f_2 & d_2 - \lambda & f_3 \\ 0 & f_3 & d_3 - \lambda \end{vmatrix} \\ &= (d_3 - \lambda) \det \begin{vmatrix} d_1 - \lambda & f_2 \\ f_2 & d_2 - \lambda \end{vmatrix} - f_3 \det \begin{vmatrix} d_1 - \lambda & f_2 \\ 0 & f_3 \end{vmatrix} \\ &= p_2(\lambda)(d_3 - \lambda) - f_3^2(d_1 - \lambda) \end{aligned}$$

For $n = r$,

$$p_r(\lambda) = p_{r-1}(\lambda)(d_r - \lambda) - f_r^2 p_{r-2}(\lambda)$$

where $p_0 = 1$. A set of polynomials $\{p_n(\lambda)\}$ is said to form a Sturm sequence.

15.2 Solution by bisection method

For $n = 1$

$$\lambda^{(1)} = d_1.$$

For $n = 2$

$$\begin{aligned}
\lambda_1^{(2)} &= 0.5(d_1 + d_2) - 0.5\left[(d_1 - d_2)^2 + 4f_2^2\right]^{1/2}, \\
\lambda_2^{(2)} &= 0.5(d_1 + d_2) + 0.5\left[(d_1 - d_2)^2 + 4f_2^2\right]^{1/2}.
\end{aligned}$$

For $d_1 > d_2$,

$$\lambda_1^{(2)} < \lambda^{(1)} < \lambda_2^{(2)}$$

(For $d_1 > d_2$, $\lambda_1^{(2)} > \lambda^{(1)} > \lambda_2^{(2)}$)

This relation can be generalized to the roots of any order of Sturm sequence

$$\lambda_1^{(r+1)} < \lambda_1^{(r)} < \lambda_2^{(r+1)} < \lambda_2^{(r)} < \lambda_3^{(r+1)} < \lambda_3^{(r)} < \lambda_4^{(r+1)} \dots$$

(1) If any one of the matrix elements f_n is zero, the matrix can be split into two separate submatrices. In this case we can solve the two submatrices separately.

(2) $p_{r+1}(\lambda)$ and $p_r(\lambda)$ cannot share a common root if all $f_n \neq 0$.

Proof: if

$$p_{r+1}(\lambda_0) = p_r(\lambda_0)$$

we can conclude all $p_n(\lambda_0) = 0$, including $p_0 = 0$, which is in contradiction with the assumpation of $p_0 = 1$.

(3) To prove the inequality relation, we need also to show that there is only one root of $p_r(\lambda)$ between any two adjacent roots of $p_{r+1}(\lambda)$.

(4) To find the lower and upper limits of all roots, we have to show

$$|\lambda_i| \leq \lambda_{\mathsf{max}}$$

where

$$\lambda_{\mathsf{max}} = \max \left\{ |f_i| + |d_i| + |f_{i+1}| \right\}.$$

(5)

$$\lim_{v \to +\infty} p_r(v) = (-v)^r$$

From this property, we conclude that ($\epsilon > 0$)

$$p_r(\lambda_{\text{max}} + \epsilon) \begin{cases} > 0 & \text{if} \quad r \quad \text{is} \quad \text{even} \\ < 0 & \text{if} \quad r \quad \text{is} \quad \text{odd} \end{cases}$$

Let us use s to record the number of agreements in sign at any given value of v between any two polynomials differing in degree by 1

(a) When $v = \lambda_{\text{max}} + \epsilon$, no two adjacent polynomial have the same sign. In this case, $s = 0$.

(b) When $v = \lambda_n^{(n)} - \epsilon$, we have $s = 1$ since $p_n(v)$ and $p_{n-1}(v)$ have the same sign at that point.

(c) When $v = \lambda_{n-1}^{(n)} - \epsilon$, we have $s = 2$

⋮

(d) When $v = \lambda_1^{(n)} - \epsilon$, we have $s = n$

15.3 QR and QL-transformation

As practical methods for diagonalizing real symmetric matrices it is most efficient to apply the QR and QL algorithm after the matrices are reduced to a tri-diagonal form using the method of Householder.

It can be shown that an symmetric matrix can always be reduced to a product of two matrices

$$A = QR$$

where Q is a unitary matrix and R is an upper triangular matrix. This is similar in spirit to the LU-decomposition of mmatrix.

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1N} \\ q_{21} & q_{22} & \cdots & q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1} & q_{N2} & \cdots & q_{NN} \end{pmatrix}$$

For a unitary matrix $q_{ij} = q_{ji}$, the number of unknowns is reduced to $N(N+1)/2$, just like a triangle matrix.

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{NN} \end{pmatrix}$$

Using the unitary property of Q, we obatin the relation

$$Q^{\dagger} A = Q^{-1} A = R.$$

Alternatively, we can define a new matrix B such that

$$B = RQ = Q^{\dagger}AQ$$

Since B is the unitary transformation of A, their eigenvalues are identical to each other. In general B is not a triangle matrix, since multiplying the triangle matrix R by Q produces nonvanishing matrix elements in the other half of the triangle. However if we define

$$A(1) = A$$

and apply a series of unitary transformations given by the recursion relation,

$$A(k+1) = Q^{\dagger}(k)A(k)Q(k)$$

for $k = 1, 2, \cdots$. It is possible to prove, when $k \to +\infty$

$$A(k) \to D \text{ (a \quad diagonal \quad matrix).}$$

As a general method for matrix diagonalization, the QL- and QR-method are not very useful because of the iterative nature of the procedure. For the tridiagonal matrices, this method is likely to be efficient. As practical methods for diagonalizing real symmetric matrices, it is possible to apply the QL- or QR-algorithm after the matrices are reduced to a tridiagonal form using, for example, the method of Householder.

In general the rate that an off-diagonal matrix element a_{ij} goes to zero is proportional to the ratio of the absolute values of two eigenvalues

$$\text{rate}\left\{a_{ij}^{(k)} \to 0\right\} \sim \left|\frac{\lambda_i}{\lambda_j}\right|^k$$

As a result, the convergence of a QL-algorithm can be slow if two eigenvalues happen to be very close to each other in their absolute values. To prevent this from happening, we can shift all diagonal values by a constant η_k. That is instead of A we work with

$$A' = A - \eta_k \mathbf{1}$$

The transformation is not affected by the shift. The ratio becomes

$$\text{rate}\left\{a_{ij}^{(k)} \to 0\right\} \sim \left|\frac{\lambda_i - \eta_k}{\lambda_j - \eta_k}\right|^k$$

If λ_i is not exactly equal to λ_j we can choose the value of η_k to be as close to that of λ_i as possible ahead. As a result the ratio can be very small.

Degeneracy: it cannot take place in a tri-diagonal matrix unless one of the superdiagonal elements vanishes.

Example:

$$c_2 = \begin{pmatrix} d_1 & f_2 \\ f_2 & d_2 \end{pmatrix}$$

15.4 The unitary matrix Q

We shall now turn our attention to the question of finding the unitary matrix Q. Consider

$$Q^\dagger(k)A(k) = L(k).$$

Each $Q^\dagger(k)$ consists of $n-1$ transformation operators $P^\dagger_m(k)$.

$$Q^\dagger(k) = P^\dagger_2(k)P^\dagger_3(k)\cdots P^\dagger_n(k)$$

For example,

$$P_n(k) = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & 0 & 0 \\ 0 & \cdots & \cos\theta^{(n)} & -\sin\theta^{(n)} \\ 0 & \cdots & \sin\theta^{(n)} & \cos\theta^{(n)} \end{pmatrix}$$

will annihilate the element $a_{n-1,n}^{(k)}$, and $P_m(k)$ will annihilate the transformed element $a\prime_{m-1,m}^{(k)}$ after $n-m$ transformations. For example, a 2×2 matrix,

$$p^\dagger = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$p^\dagger \begin{pmatrix} d_1 & f_2 \\ f_2 & d_2 \end{pmatrix} = \begin{pmatrix} d_1c - f_2s & f_2c - d_2s \\ f_2c + d_1s & d_2c + f_2s \end{pmatrix}$$

Take $f_2c - d_2s = 0$,

$$\tan\theta = \frac{f_2}{d_2}.$$

To evaluate the eigenvalues and eigenvectors of a real symmetric matrix,

(1) to a tri-diagonal form by using the Householder method

$$R^T A R = J$$

(2) to diagonal form by using the QL- or QR-transformation.

$$Q^T J Q = D$$

(3) Thus, $A \rightarrow D$,

$$V^T A V = D$$

where $V = RQ$.

(4). Eigenvector v_m with eigenvalue d_m,

$$v_m = \begin{pmatrix} v_{1m} \\ v_{2m} \\ \vdots \\ v_{nm} \end{pmatrix}$$

15.5 Complex Matrices

Since the Hamiltonian is Hermitian, in most cases, it is complex, not just real. Fortunately, the problem can be reduced to a real symmetric one. For example, a complex C can be written as a sum of two real matrices A and B

$$C = A + iB$$

Let $v = x+iy$ (x and y are real) be an eigenvector of C with an eigenvalue λ

$$Cv = \lambda v$$

$$(A + iB)(x + iy) = \lambda(x + iy)$$

$$\begin{aligned} Ax - By &= \lambda x \\ Ay + Bx &= \lambda y \end{aligned}$$

It is noticed that for a Hermitian matrix its eigenvalues are real. In a matrix form,

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

$$C'v' = \lambda v'$$

In other words, the complex eigenvalue problem is equivalent to a problem of real matrix. If C is Hermitian,

$$c_{ij}^* = c_{ji}$$

the elements in A and B have

$$\begin{aligned} a_{kl} &= a_{lk} \\ b_{kl} &= -b_{lk} \end{aligned}$$

As a result, $C\prime$ is a real, symmetric matrix. In this way, the eigenvalue problem for a complex, but Hermitian matrix reduces to one for a real, symmetric matrix of twice the dimension.

16 The Lanczos Method

For the eigenvalue problems it is possible to use the Lanczos method to construct a tridiagonal basis for the Hamiltonian operator. There are two advantages in using a tridiagonal basis. The first is the obvious case in diagonalizing such a matrix. The second is in the amount of computer memory required to store the matrix.

Let us start with the Schrodinger equation. Consider an arbitrary physical state or vector matrix, $\left|\Phi_1\right\rangle$, and the state should be normalized,

$$\begin{aligned}\left|\phi_1\right\rangle &= \left|\Phi_1\right\rangle/(\left\langle\Phi_1\right.\left|\Phi_1\right\rangle)^{1/2}\\ d_1 &= \left\langle\phi_1\right| H \left|\phi_1\right\rangle\end{aligned}$$

The projection operator:

$$\begin{aligned} P_1 &= |\phi_1\rangle \langle\phi_1| \\ P_1^2 &= P_1 \end{aligned}$$

The physical meaning of a projection operator is that it project out the component of the state in an arbitary state

$$\begin{aligned} P_1 |\Psi\rangle &= |\phi_1\rangle \langle\phi_1| \Psi\rangle \propto |\phi_1\rangle \\ P_1 |\phi_1\rangle &= |\phi_1\rangle \\ (1 - P_1) |\phi_1\rangle &= 0 \end{aligned}$$

The state, $|\Phi_2\rangle$, is constructed to be normal to the state $|\phi_1\rangle$,

$$\begin{aligned} |\Phi_2\rangle &= (1 - P_1) H |\phi_1\rangle \\ |\phi_2\rangle &= |\Phi_2\rangle /(\langle\Phi_2 |\Phi_2\rangle)^{1/2} \rightarrow \langle\phi_2|\phi_1\rangle = 0 \end{aligned}$$

$$\begin{aligned} H\,|\phi_1\rangle &= c_1\,|\phi_1\rangle + c_2\,|\phi_2\rangle \\ [P_1 + (1-P_1)H]\,|\phi_1\rangle &= P_1 H\,|\phi_1\rangle + (1-P_1)H\,|\phi_1\rangle \\ &= c_1\,|\phi_1\rangle + c_2\,|\phi_2\rangle \\ d_2 &= \langle\phi_2|\,H\,|\phi_2\rangle \\ f_2 &= \langle\phi_1|\,H\,|\phi_2\rangle \end{aligned}$$

If $|\phi_1\rangle$ is the eigenstate of H, then $|\Phi_2\rangle = 0$. Generally $|\phi_1\rangle$ is not an eigenstate. The newly constructed state $|\phi_2\rangle$ is othorgonal to $|\phi_1\rangle$, that is, $\langle\phi_2|\phi_1\rangle = 0$.

The state, $|\Phi_3\rangle$, othorgonal to both $|\phi_1\rangle$ and $|\phi_2\rangle$

$$\begin{aligned} |\Phi_3\rangle &= (1-P_1-P_2)H\,|\phi_2\rangle \\ |\phi_3\rangle &= |\Phi_3\rangle\,/(\langle\Phi_3\,|\Phi_3\rangle)^{1/2} \\ H\,|\phi_2\rangle &= c_1\,|\phi_1\rangle + c_2\,|\phi_2\rangle + c_3\,|\phi_3\rangle \\ d_3 &= \langle\phi_3|\,H\,|\phi_3\rangle \\ f_3 &= \langle\phi_2|\,H\,|\phi_3\rangle \end{aligned}$$

$$\begin{aligned}(1 - P_1 - P_2)\left|\phi_1\right\rangle &= 0 \\ \left\langle\phi_1\right| H \left|\phi_3\right\rangle &= 0 \\ H\left|\phi_1\right\rangle &= c_1\left|\phi_1\right\rangle + c_2\left|\phi_2\right\rangle\end{aligned}$$

The state, $\left|\Phi_4\right\rangle$

$$\begin{aligned}\left|\Phi_4\right\rangle &= (1 - P_2 - P_3)H\left|\phi_3\right\rangle \\ \left|\phi_4\right\rangle &= \left|\Phi_4\right\rangle/(\left\langle\Phi_4\right|\left.\Phi_4\right\rangle)^{1/2} \\ H\left|\phi_3\right\rangle &= c_2\left|\phi_2\right\rangle + c_3\left|\phi_3\right\rangle + c_4\left|\phi_4\right\rangle \\ d_4 &= \left\langle\phi_4\right| H\left|\phi_4\right\rangle \\ f_4 &= \left\langle\phi_3\right| H\left|\phi_4\right\rangle\end{aligned}$$

$$\left\langle\phi_i\right| H\left|\phi_4\right\rangle = 0, \quad \text{for} \quad i = 1, 2$$

In a general form,

$$H\left|\phi_n\right\rangle = c_{n-1}\left|\phi_{n-1}\right\rangle + c_n\left|\phi_n\right\rangle + c_{n+1}\left|\phi_{n+1}\right\rangle$$

On the set of base state $\{|\phi_i\rangle\}$, the Hamiltonian is reduced to a tri-diagonal matrix form.

$$H = \begin{pmatrix} d_1 & f_2 & 0 & \cdots & 0 \\ f_2 & d_2 & f_3 & \cdots & 0 \\ 0 & f_3 & d_3 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & f_n \\ 0 & \cdots & \cdots & f_n & d_n \end{pmatrix}$$

16.0.1 A simple method to find the lowest energy state

A simple method to find the lowest energy state: starting with any state $|\phi_1\rangle$ we can construct a state $|\phi_2\rangle$ normal to the state $|\phi_1\rangle$. On these two states, the Hamiltonian is written as

$$H' = \begin{pmatrix} d_1 & f_2 \\ f_2 & d_2 \end{pmatrix}.$$

Its two eigenvalues are

$$\begin{aligned} \lambda_1 &= 0.5(d_1 + d_2) - 0.5\left[(d_1 - d_2)^2 + 4f_2^2\right]^{1/2}, \\ \lambda_2 &= 0.5(d_1 + d_2) + 0.5\left[(d_1 - d_2)^2 + 4f_2^2\right]^{1/2}. \end{aligned}$$

The first one has a lower energy than either d_1 or d_2. The corresponding eigenstate is

$$H' \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

$$d_1 c_1 + f_2 c_2 = \lambda_1 c_1$$

$$
\begin{aligned}
c_1 &= \frac{\lambda_1 - f_2}{\left(d_1^2 + (\lambda_1 - f_2)^2\right)^{1/2}} \\
c_2 &= \frac{d_1}{\left(d_1^2 + (\lambda_1 - f_2)^2\right)^{1/2}}
\end{aligned}
$$

Take the new initial state

$$
|\phi_1\rangle \to c_1 |\phi_1\rangle + c_2 |\phi_2\rangle
$$

Proceeding to repeat the process, we have

$$
\begin{aligned}
|\phi_1\rangle &\to \text{the true ground state} \\
\lambda_1 &\to \text{the ground state energy of H}
\end{aligned}
$$

16.1 Example: Anharmonic oscillator

$$H = H_0 + V$$

$$H_0 = -\frac{\hbar^2}{2\mu}\frac{d^2}{dx^2} + \frac{1}{2}\mu\omega^2 x^2$$

$$V = \epsilon\hbar\omega\left(\frac{\mu\omega}{\hbar}\right)^2 x^4$$

The ground state energy can be calculated by means of Lanczos method.

17 Exact Diagonalization Study of Antiferromagnetism in the Heisenberg Model

17.1 The model

The model Hamiltonian for the antiferromagnetic Heisenberg model is

$$H = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j$$

where the sum runs over the nearest neighbor sites. We only consider the case of spin $S = 1/2$. The model is defined in a lattice. The common lattices we study are

- One-dimensional chain

- Ladder lattice

- Two-dimensional square lattice, et al.

To decrease the finite-size effect, we often take a periodic boundary condition. The open boundary condition and the anti-periodic boundary condition are also used.

17.2 Basic properties for Spin S=1/2 operators

Define the increasing and decreasing operators

$$\begin{aligned}\mathbf{S}_i^+ &= \mathbf{S}_i^x + i\mathbf{S}_i^y \\ \mathbf{S}_i^- &= \mathbf{S}_i^x - i\mathbf{S}_i^y\end{aligned}$$

Denote by $|\uparrow\rangle$ and $|\downarrow\rangle$ the two eigenstates of $\mathbf{S}_i^z$, respectively. We have

$$\begin{aligned}\mathbf{S}_i^z\left|\uparrow\right\rangle &= \frac{\hbar}{2}\left|\uparrow\right\rangle \\ \mathbf{S}_i^z\left|\downarrow\right\rangle &= -\frac{\hbar}{2}\left|\downarrow\right\rangle\end{aligned}$$

and

$$\begin{aligned}\mathbf{S}_i^+\left|\uparrow\right\rangle &= 0, \quad \mathbf{S}_i^+\left|\downarrow\right\rangle = \hbar\left|\uparrow\right\rangle \\ \mathbf{S}_i^-\left|\uparrow\right\rangle &= \hbar\left|\downarrow\right\rangle, \quad \mathbf{S}_i^-\left|\downarrow\right\rangle = 0\end{aligned}$$

We take $\hbar$ as the unit of the angular momentum.

Two good quantum numbers

(1). The total spins, $\mathbf{S}_{tot} = \sum_i \mathbf{S}_i$

$$[\mathbf{S}_{tot}^2, H] = 0$$

(2). Its z-component

$$[\mathbf{S}_{tot}^z, H] = 0$$

Therefore we can diagonalize the Hamiltonian in block according to S_{tot} and S_{tot}^z. Usually, it is not easy to determine S_{tot} in practice, but easy to determine S_{tot}^z by the difference of the numbers of spin-up and spin-down. In most cases, we restrict ourselves in a certain S_{tot}^z, which can decrease the dimension of the matrix drastically. Since the system processes the SU(2) symmetry, each state with S_{tot} has other $2S_{tot}$ degenerate states with the same total spin and different S^z_{tot}, we just consider the subspace with $S^z_{tot} = 0$.

17.3 Two-site problem and the exchange operator

Denote by $|m_1 m_2\rangle$ the base kets for a two-site system. There are four possible configurations.

$$|\uparrow\uparrow\rangle\,, \;|\uparrow\downarrow\rangle\,, \;|\downarrow\uparrow\rangle\,, \;|\downarrow\downarrow\rangle\,.$$

The two-site Hamiltonian

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2}\left(\mathbf{S}_1^+ \cdot \mathbf{S}_2^- + \mathbf{S}_1^- \cdot \mathbf{S}_2^+\right) + \mathbf{S}_1^z \cdot \mathbf{S}_2^z$$

acts on these base kets, and we have

$$\begin{aligned}
\mathbf{S}_1\cdot\mathbf{S}_2\,|\uparrow\uparrow\rangle &= \frac{1}{4}\,|\uparrow\uparrow\rangle \\
\mathbf{S}_1\cdot\mathbf{S}_2\,|\uparrow\downarrow\rangle &= \frac{1}{2}\,|\downarrow\uparrow\rangle - \frac{1}{4}\,|\uparrow\downarrow\rangle \\
\mathbf{S}_1\cdot\mathbf{S}_2\,|\downarrow\uparrow\rangle &= \frac{1}{2}\,|\uparrow\downarrow\rangle - \frac{1}{4}\,|\downarrow\uparrow\rangle \\
\mathbf{S}_1\cdot\mathbf{S}_2\,|\downarrow\downarrow\rangle &= \frac{1}{4}\,|\downarrow\downarrow\rangle
\end{aligned}$$

These four equation can be expressed in a more compact form

$$\left(2\mathbf{S}_1\cdot\mathbf{S}_2+\frac{1}{2}\right)|m_1m_2\rangle = |m_2m_1\rangle$$

Based on the equation, we define an exchange operator

$$P_{ij} = 2\mathbf{S}_i\cdot\mathbf{S}_j+\frac{1}{2}$$

which exchange the two states on the site i and j. In terms of the exchange operator, the Hamiltonian is rewritten as

$$H = \frac{J}{2} \sum P_{ij}$$

We have removed a constant which does not affect the physical properties of the system.

17.4 Base Ket

To study this system by exact diagonalization method, we first have to introduce a complete and orthogonal set of base kets. Since each site has two possible states $|\uparrow\rangle$ and $|\downarrow\rangle$, we can use 1 and 0 to represent the two states. For a many body system, we can introduce a binary to represent a possible configuration. For example,

$N=2$	Binary	Code
$\lvert\uparrow\uparrow\rangle$	11	1
$\lvert\uparrow\downarrow\rangle$	10	2
$\lvert\downarrow\uparrow\rangle$	01	3
$\lvert\downarrow\downarrow\rangle$	00	4

When an exchange operator acts on one basis, for example 10, we get new number 01. In the two site problem, $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are two eigenstates of energy. Linear combination of other two states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ leads to two eigenstates

$$\begin{aligned} P_{12}\,|\uparrow\downarrow\rangle &= |\downarrow\uparrow\rangle\text{, i.e., } 10\rightarrow 01 \\ P_{12}\,|\downarrow\uparrow\rangle &= |\uparrow\downarrow\rangle\text{, i.e., } 01\rightarrow 10 \end{aligned}$$

$$\begin{aligned} P_{12}\left(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle\right) &= \left(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle\right) \\ P_{12}\left(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle\right) &= -\left(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle\right) \end{aligned}$$

These two states have $S^z_{tot} = 0$ and $S_{tot} = 1$ and 0, respectively.

For $N = 4$, i.e., a four-site problem, there are 16 possible configurations. As the S^z_{tot} is a good quantum number, we just consider

$N = 4$	Binary	Code
$\lvert\uparrow\uparrow\downarrow\downarrow\rangle$	1100	12→1
$\lvert\uparrow\downarrow\uparrow\downarrow\rangle$	1010	10→2
$\lvert\uparrow\downarrow\downarrow\uparrow\rangle$	1001	9→3
$\lvert\downarrow\uparrow\uparrow\downarrow\rangle$	0110	6→4
$\lvert\downarrow\uparrow\downarrow\uparrow\rangle$	0101	5→5
$\lvert\downarrow\downarrow\uparrow\uparrow\rangle$	0011	3→6

Let us see an example,

$$\begin{aligned} P_{23}\,\lvert\uparrow\downarrow\uparrow\downarrow\rangle &= \lvert\uparrow\uparrow\downarrow\downarrow\rangle \\ 1010 &\rightarrow 1100 \\ 2 &\rightarrow 1 \end{aligned}$$

Four-site problem can be solved analytically. Assume the four spins form a ring.

$$H = P_{12} + P_{23} + P_{34} + P_{41}$$

In the subspace $S^z_{tot} = 0$, the problem is reduced to an eigensystem of a 6×6 , not 16×16 matrix.

Without loss of the generality, we consider a system with N sites (N is even). We introduce a binary number to represent the state. Suppose the Mth base ket is written as

$$base[M] = m_1 m_2 \cdots m_i \cdots m_j \cdots m_M$$

where $m = 0$ or 1. For $\mathrm{S}_{tot} = 0$, we constraint

$$\sum_i m_i = \frac{N}{2}$$

The dimension of the base ket is

$$d_{\mathsf{max}} = C_N^{N/2} = \frac{N!}{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!}$$

Therefore M can be from 1 to d_{max}. When P$_{ij}$ acts on base[M], we have a new base

$$base[M_{ij}] = m_1 m_2 \cdots m_j \cdots m_i \cdots m_M$$

Notice that only ith and jth numbers exchange their positions. This is a new base ket. The new binary number gives a new code for the base ket and we should evaluate the code M_{ij}.

Any state with $S_{totZ} = 0$ can be expanded in terms of $\{base[M]\}$. We can use a d_{max}-dimensional vector to represent it.

$$|\phi\rangle = \left\{c_1, c_2, \cdots c_{d_{\mathsf{max}}}\right\}$$

When an exchange operator acts on it, we have a new state

$$P_{ij}\left|\phi\right\rangle = \left|\varphi\right\rangle = \left\{f_1, f_2, \cdots f_{d_{\mathsf{max}}}\right\}$$

The relation between $|\phi\rangle$ and $|\varphi\rangle$ is

$$\begin{aligned} c_n &\rightarrow f_m \\ base[n] &\rightarrow base[m] \end{aligned}$$

under the operation of P_{ij}

Normalization of a state

$$\begin{aligned} \langle\phi|\phi\rangle &= \sum_i c_i^2 \\ c_n &\rightarrow c_n \left(\sum_i c_i^2\right)^{-1/2} \end{aligned}$$

17.5 Lanczos Method

We are now in the position to evaluate the ground state energy and ground state by utilizing the Lanczos method.

(1) Starting from the state $|\phi_0\rangle$, for example, $\{1, 1, \cdots, 1\}$. we first normalize the state $\left|\tilde{\phi}_0\right\rangle$ such that

$$\left\langle\tilde{\phi}_0|\tilde{\phi}_0\right\rangle = 1$$

The average value of H on the state $\left|\tilde{\phi}_0\right\rangle$ is

$$a_0 = \left\langle\tilde{\phi}_0|H|\tilde{\phi}_0\right\rangle$$

(2) H acts on the state $\left|\tilde{\phi}_0\right\rangle$. We can construct a state orthogonal to the state $\left|\tilde{\phi}_0\right\rangle$

$$\begin{aligned} |\phi_1\rangle &= \left(1 - \left|\tilde{\phi}_0\right\rangle \left\langle\tilde{\phi}_0\right|\right) H \left|\tilde{\phi}_0\right\rangle = H \left|\tilde{\phi}_0\right\rangle - a_0 \left|\tilde{\phi}_0\right\rangle \\ |\phi_1\rangle &\to \left|\tilde{\phi}_1\right\rangle \end{aligned}$$

We can evaluate

$$\begin{aligned} b_1 &= \left\langle\tilde{\phi}_0\right| H \left|\tilde{\phi}_1\right\rangle \\ a_1 &= \left\langle\tilde{\phi}_1\right| H \left|\tilde{\phi}_1\right\rangle \end{aligned}$$

(3) H acts on the state $\left|\tilde{\phi}_1\right\rangle$. We can construct a state orthogonal to the state $\left|\tilde{\phi}_0\right\rangle$ and $\left|\tilde{\phi}_1\right\rangle$

$$|\phi_2\rangle = \left(1 - \left|\tilde{\phi}_0\right\rangle \left\langle\tilde{\phi}_0\right| - \left|\tilde{\phi}_1\right\rangle \left\langle\tilde{\phi}_1\right|\right) H \left|\tilde{\phi}_1\right\rangle$$

$$\begin{aligned} &= H\left|\tilde{\phi}_1\right\rangle - b_1\left|\tilde{\phi}_0\right\rangle - a_1\left|\tilde{\phi}_1\right\rangle \\ |\phi_1\rangle &\rightarrow \left|\tilde{\phi}_1\right\rangle \end{aligned}$$

We can evaluate

$$\begin{aligned} b_2 &= \left\langle\tilde{\phi}_1\right| H \left|\tilde{\phi}_2\right\rangle \\ a_2 &= \left\langle\tilde{\phi}_2\right| H \left|\tilde{\phi}_2\right\rangle \end{aligned}$$

(4) To proceed the procedure we can construct the series of states $\left\{\left|\tilde{\phi}_n\right\rangle\right\}$ until n=d$_{\mathrm{max}}$. Then we have a tri-diagonal matrix, which can be diagonalized by QR- and QL-decomposition. However the dimension of the matrix may be very huge for many body system, instead of constructing a complete tri-diagonal matrix for the system, we can truncate the series at any n. n can be even 2. In this case several itineration are necessary. We first diagonalize the truncated matrix and evaluate the lowest energy

state $\left|\tilde{\phi}_g\right\rangle$ and the lowest energy E$_g$, which have a lower energy than the initial state $\left|\tilde{\phi}_0\right\rangle$ and any other states constructed by Lanczos method. Next we use the evaluated $\left|\tilde{\phi}_g\right\rangle$ as a new initial state $\left|\tilde{\phi}_0\right\rangle$, and repeat the same procedure to construct the tri-diagonal matrix and to evaluate the new lowest energy state and its lowest energy. Comparing the new lowest energy with the old one, if they are equal within the tolerance or error ϵ, we get the ground state and its energy. Otherwise we have to proceed the procedure until we get the lowest energy within tolerance.

(5) The evaluated ground state can be used to evaluate other properties of the system, or example, the spin-spin correlation function.

(6) We release the constraint on the form of the Hamiltonian. The coupling J may not be a constant, and can even be a disorder value. The lattice of the system is limited to a small size, about 30 - 50.

Problems

1. Solve the linear equations

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

   by using the Gauss-Jordan elimination method, and LU-decomposition method.

2. Construct all the polynomial in the Sturm sequence for the characteristic equation of the following tri-diagonal matrix:

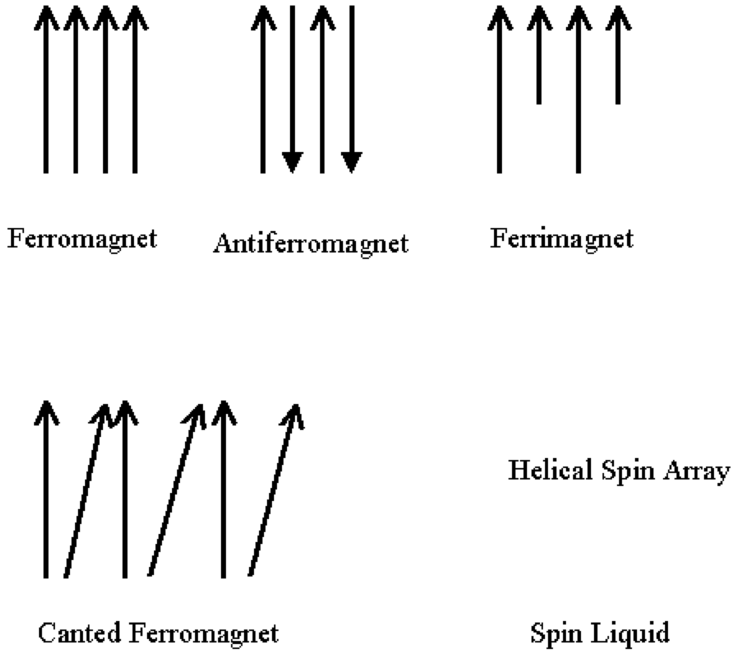
$$A = \begin{pmatrix} 10 & 5 & 0 \\ 5 & -1 & -1 \\ 0 & -1 & 10 \end{pmatrix}$$

Find the eigenvalues using the bisection method.

18 Appendix: Quantum Magnetism

Magnetism is inseparable from quantum mechanics, for a strictly classical system in thermal equilibrium can display no magnetic momentum, even in a magnetic field. The magnetic momentum of a free atom has three principal sources: the spin with which electrons are endowed, their orbital angular momentum about the nucleus; and the change in the orbital moment induced by an applied magnetic field.

Ordered arrangements of electron spins



Ferromagnet

Antiferromagnet

Ferrimagnet

Helical Spin Array

Canted Ferromagnet

Spin Liquid

18.1 Spin Exchange

Ferromagnetism is obtained in solids when the magnetic moments of many electrons align. Antiferromagnetism and spin density waves describe oscillatory ordering of magnetic moments. The classical dipolar interaction between the electron moments (which is of order $10^5 eV$) is for too weak to explain the observed magnetic transition temperature (which are of order $10^2 - 10^3\ ^0k$ in transition metal and rare earth compounds)

The coupling mechanism that gives rise to magnetism derives from the following fundamental properties of electrons:

- The electron's spin

- The electron's kinetic energy

- Pauli exclusion principle

- Coulomb repulsion

Before we introduce the physical origin of the magnetic coupling between electrons in solids, we simply review some standard definitions and basic relations of second quantization.

For an orthonormal single-particle basis, $\{|\phi_i\rangle\}$

$$\left\langle \phi_i | \phi_j \right\rangle = \delta_{ij}.$$

The creation operator of state i is $a_i^\dagger$ and its Hermitian conjugate is annihilation operator a_i. Both are defined with respect to the vacuum state $|0\rangle_i$such that

$$\begin{aligned} |\phi_i\rangle &= a_i^\dagger \, |0\rangle_i \\ a_i \, |0\rangle_i &= 0. \end{aligned}$$

The number operator is defined as

$$n_i = a_i^\dagger a_i$$

For bosons,

$$\begin{aligned}\left[a_i, a_j^\dagger\right] &= a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij} \\ \{a_i, a_j\} &= 0\end{aligned}$$

For electrons with spin $s = 1/2$, we have to introduce a pair of operators, $c_{i,\sigma}^\dagger$ where $\sigma = \uparrow, \downarrow$ such that

$$\begin{aligned}\{c_{i\sigma}, c_{j\sigma}^\dagger\} &= \delta_{ij}\delta_{\sigma\sigma,} \\ \{c_{i\sigma}, c_{j\sigma,}\} &= 0\end{aligned}$$

The spin operator can be expressed as

$$
\begin{aligned}
S_i^{\dagger} &= S_{ix} + iS_{iy} = \hbar c_{i\uparrow}^{\dagger} c_{i\downarrow} \\
S_i^{-} &= S_{ix} - iS_{iy} = \hbar c_{i\downarrow}^{\dagger} c_{i\uparrow} \\
S_i^{z} &= \frac{\hbar}{2}(c_{i\uparrow}^{\dagger} c_{i\uparrow} - c_{i\downarrow}^{\dagger} c_{i\downarrow})
\end{aligned}
$$

The commutation relations:

$$
\begin{aligned}
[S_i^{\dagger}, S_i^{-}] &= 2\hbar S_i^{z} \\
[S_i^{z}, S_i^{\pm}] &= \pm\hbar S_i^{\pm}
\end{aligned}
$$

18.2 Two-Site Problem

The Hersenberg spin exchange interaction is written as

$$H = J\mathbf{S}_1 \cdot \mathbf{S}_2$$

To obtain the eigenstates of H, we check the following relations,

(1) $[\mathbf{S}_1^2, H] = 0$,but $[\mathbf{S}_1, H] \neq 0$

(2) $[\mathbf{S}_2^2, H] = 0$,but $[\mathbf{S}_2, H] \neq 0$

(3) $[\mathbf{S}_1 + \mathbf{S}_2, H] = 0$

Therefore $\mathbf{S}_1^2$, $\mathbf{S}_2^2$,and $\mathbf{S}_{tot}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2$ and its z-component $\mathbf{S}_{tot}^z$ are good quantum numbers, but S_1^z, S_2^z are not. Hence we can denote the simultaneous eigenkets of $\mathbf{S}_{tot}^2$, $\mathbf{S}_{tot}^z$, $\mathbf{S}_1^2$ and $\mathbf{S}_2^2$ by

$$|S_{tot}, S_{tot}^z, S_1, S_2\rangle$$

such that

$$\begin{aligned}
\mathbf{S}_1^2 |S_{tot}, S_{tot}^z, S_1, S_2\rangle &= S_1(S_1 + 1) |S_{tot}, S_{tot}^z, S_1, S_2\rangle \\
\mathbf{S}_2^2 |S_{tot}, S_{tot}^z, S_1, S_2\rangle &= S_2(S_2 + 1) |S_{tot}, S_{tot}^z, S_1, S_2\rangle \\
\mathbf{S}_{tot1}^z |S_{tot}, S_{tot}^z, S_1, S_2\rangle &= S_{tot}(S_{tot} + 1) |S_{tot}, S_{tot}^z, S_1, S_2\rangle \\
\mathbf{S}_{tot1}^z |S_{tot}, S_{tot}^z, S_1, S_2\rangle &= S_{tot}^z |S_{tot}, S_{tot}^z, S_1, S_2\rangle
\end{aligned}$$

Fortunately, the state kets are also the eigenkets of H:

$$H\left|S_{tot}, S_{tot}^{z}, S_1, S_2\right\rangle = E\left|S_{tot}, S_{tot}^{z}, S_1, S_2\right\rangle$$

where

$$E = \frac{J}{2}\left[S_{tot}(S_{tot}+1) - S_1(S_1+1) - S_2(S_2+1)\right]$$

and

$$S_{tot} = \left|S_1 - S_2\right|, \left|S_1 - S_2\right| + 1, ...S_1 + S_2$$

Since the energy eigenvalues are independent of S_{tot}^{z} can be $-S_{tot}, ...S_{tot}$ the energy eigenstates are $(2S_{tot}+1)-$fold degenerated. From the point of view of symmetry, the degeneracy of the eigenstates originates from the invariance of H under the $SU(2)$ symmetry rotation,

$$UHU^{\dagger} = H$$

where

$$U = \exp[-i\mathbf{S}_{tot} \cdot \mathbf{n}\phi/\hbar]$$

The ground state:

The lowest energy state is determined by the sign of J: when $J > 0$, S_{tot} should be taken to be minimum, otherwise S_{tot} should be taken to be maximal.

The case of $J > 0$:

$$S_{tot} = S_1 - S_2,\ (S_1 > S_2)$$

The two spins are antiparallel, which is called antiferromagnetic

The ground state energy

$$E_g = -J(S_1 + 1)S_2$$

The case of $J < 0$

$$S_{tot} = S_1 + S_2$$

The two spins are parallel, which is ferromagnetic

$$E_g = -\left|J\right| S_1 S_2$$

18.3 Ferromagnetic Exchange $(J < 0)$

Ferromagnetic exchange coupling originates from the direct Coulomb interaction and the Pauli exclusion principle. In the second quantized form, the two-body Coulomb interaction is given by

$$V = \frac{1}{2}\int dxdy\tilde{v}(x,y)\mathbf{\Psi}_s^\dagger(x)\mathbf{\Psi}_{s'}^\dagger(y)\mathbf{\Psi}_{s'}(y)\mathbf{\Psi}_s(x)$$

The field operator

$$\mathbf{\Psi}_s^\dagger(x) = \sum_i \phi_i^*(x)c_{is}^\dagger.$$

The interaction can be expressed as

$$V \ = \ \frac{1}{2}\int dxdy\tilde{v}(x,y)\phi_i^\dagger(x)\phi_j^\dagger(y)\phi_k(y)\phi_l(x)c_{is}^\dagger c_{js'}^\dagger c_{ks'}c_{ls}$$

$$
\begin{aligned}
= \; & \sum_i U_{ii} n_{is} n_{i-s} + \sum_{i,i'} U_{ii'} (n_{i\uparrow} + n_{i\downarrow})(n_{i'\uparrow} + n_{i'\downarrow}) \\
& + \sum_{i,i'} J^F_{ii'} c^\dagger_{is} c^\dagger_{i's'} c_{is'} c_{i's} + \cdots
\end{aligned}
$$

where

$$
\begin{aligned}
U_{ii'} &= \frac{1}{2} \int dxdy \, |\phi_i(x)|^2 \, |\phi_{i'}(y)|^2 \, \tilde{v}(x,y) \\
J^F_{ii'} &= \frac{1}{2} \int dxdy \tilde{v}(x,y) \phi^\dagger_{i'}(x) \phi^\dagger_i(y) \phi_{i'}(y) \phi_i(x)
\end{aligned}
$$

The exchange interaction J^Facts as a Hersenberg interaction:

$$
J^F_{ij'} \sum c^+_{is} c^+_{i's'} c_{is'} c_{i's} = -2 J^F_{ij'} \sum (S_i \cdot S_{i'} + \frac{1}{4} n_i \cdot n_{i'})
$$

The positivity of $J^F_{ij'}$ can be proved as follows:

(1) Complete screening:

$$v = \delta(x - y)$$

In the case,

$$J^F_{ii'} = \frac{1}{2} \int dx \, |\phi_i(x)|^2 \, |\phi_{i'}(x)|^2 > 0$$

(2) Long-ranged Coulomb interaction

$$\tilde{v} = \frac{e^2}{|\vec{x} - \vec{y}|}$$

Assume ϕ_i is the plane wave

$$\begin{aligned} J^F_{ii'} &\propto \int dx\, \mathsf{exp}[ik\cdot x]\frac{e^2}{|x|} \\ &= 4\pi\frac{e^2}{k^2} > 0 \end{aligned}$$

This is ferromagnetic!

18.4 Antiferromagnetic Exchange

Two-site (or atom) problem with two electrons

Let's consider two orthogonal orbitals localized on two atoms labelled by $i = 1, 2$. Tunnelling between the two atoms (or states) is described by a hopping Hamiltonian

$$H^t = -t \sum_s (C_{1s}^+ C_{2s} + C_{2s}^+ \, C_{1s})$$

For simplicity, we consider an on-site interaction

$$H_u = U \sum n_{i\uparrow} \cdot n_{i\downarrow}$$

To explore the physical origin of anti-ferromagnetic coupling, we consider a special case: $u >> t$. In the case, we choose H_u to be the zero-order(or unperturbed) Hamiltonian and H^+ to be the perturbation.

For $H_{u:}$ there are six possible configurations which are eigenstates of H$_u$

(1) E=0, $|1\uparrow,2\uparrow\rangle$, $|1\uparrow,2\downarrow\rangle$, $|1\downarrow,2\uparrow\rangle$, $|1\downarrow,2\downarrow\rangle$;

(2) E=U, $|1\uparrow,1\downarrow\rangle$, $|2\downarrow,2\uparrow\rangle$

Denote $|\alpha\rangle$ the unperturbed state with energy 0 and $|n\rangle$ denote the two state with E=U (n=1,2). In terms of the c-operators,

$$|\alpha\rangle = c^{+}_{1s}c^{+}_{2s'}|0\rangle$$

and

$$\begin{aligned} |1\rangle &= C^{\dagger}_{1\uparrow}C^{\dagger}_{1\downarrow}|0\rangle \\ |2\rangle &= C^{\dagger}_{2\uparrow}C^{\dagger}_{2\downarrow}|0\rangle \end{aligned}$$

In the first-order perturbation theory

$$\left\langle \alpha \left| H_t \right| \alpha' \right\rangle = 0$$

In the second-order perturbation theory

$$\begin{aligned} \left\langle \alpha \left| \Delta H^{(2)} \right| \alpha \right\rangle &= \sum_{n=1,2} \frac{\left\langle \alpha \left| H^t \right| n \right\rangle \left\langle n \left| H^t \right| \alpha \right\rangle}{\langle \alpha \left| H_u \right| \alpha \rangle - \langle n \left| H_u \right| n \rangle} \\ &= -\frac{1}{U} \sum \left\langle \alpha \left| H^t \right| n \right\rangle \left\langle n \left| H^t \right| \alpha \right\rangle \end{aligned}$$

$|n\rangle \langle n|$ is a projection operator

$$|1\rangle \langle 1| \;\; = \;\; n_{1\uparrow} n_{1\downarrow} (1 - n_{2\uparrow})(1 - n_{2\downarrow})(\uparrow\downarrow -)$$

$$|2\rangle\langle 2| \;=\; (1-n_{1\uparrow})(1-n_{1\downarrow})n_{2\uparrow}n_{2\downarrow}(-\uparrow\downarrow)$$

$$\begin{aligned}
& \left\langle \alpha \left| H^t \right| 1 \right\rangle \left\langle 1 \left| H^t \right| \alpha' \right\rangle \\
= & \langle\alpha| (C^\dagger_{2\uparrow}C_{\dot{1}\uparrow} + C^\dagger_{2\downarrow}C_{\dot{1}\downarrow} + \hbar.c)n_{1\uparrow}n_{1\downarrow}(1-n_{2\uparrow})(1-n_{2\downarrow})(C^\dagger_{2\uparrow}C_{\dot{1}\uparrow} + C^\dagger_{2\downarrow}C_{\dot{1}\downarrow} + \hbar.c) \Big| \\
= & \langle\alpha| (C^\dagger_{2\uparrow}C_{\dot{1}\uparrow} + C^\dagger_{2\downarrow}C_{\dot{1}\downarrow})n_{1\uparrow}n_{1\downarrow}(1-n_{2\uparrow})(1-n_{2\downarrow})(C^\dagger_{1\uparrow}C_{\dot{2}\uparrow} + C^\dagger_{1\downarrow}C_{\dot{2}\downarrow}) \left|\alpha'\right\rangle \\
= & \langle\alpha| (-C^\dagger_{1\downarrow}C_{\dot{1}\uparrow}C^\dagger_{2\uparrow}C_{\dot{1}\downarrow} - C^\dagger_{1\uparrow}C_{\dot{1}\downarrow}C^\dagger_{2\downarrow}C_{\dot{2}\uparrow}) + n_{1\uparrow}n_{2\downarrow} + n_{1\downarrow}n_{2\uparrow}) \left|\alpha'\right\rangle \\
= & \langle\alpha| - 2S_1 \cdot S_2 + \frac{1}{2}(n_{1\uparrow} + n_{1\downarrow})(n_{2\uparrow} + n_{2\downarrow}) \left|\alpha'\right\rangle
\end{aligned}$$

Therefore the effective Hamiltonian $\Delta H^{(2)}$ can be written as an isotropic antiferromagnetic Hersenberg spin exchange form.

$$\Delta H^{(2)} = +\frac{4t^2}{U}(\vec{S}_1 \cdot \vec{S}_2 - \frac{1}{4}\ )$$

As $4t^2/U > 0$, the exchange coupling is antiferromagnetic! The ground state of this two-site problem is spin singlet, i.e. $S = 0$. Our discussion on the two-site problem can be easily generalized to a many-site system. The Hersenberg model is defined on a lattice

$$H = \sum J_{ij} S_i \cdot S_j$$

where i and j are the lattice sites and usually are of the nearest neighbour pair. S_i can be taken any value of half-integer, $S = 1/2, 1 \ldots$

One of the most important application of spin superexchange with current interests is the high temperature superconductivity. the so-called "$t - J$" model is extensively discussed over the last decade.

$$H = -t \sum_{\langle ij \rangle, \sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + \frac{4t^2}{U} \sum_{i,j} (\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j)$$

which is limited within the Hilbert space excluding double occupancy of electrons on the same site.