Теория возмущений в линейной алгебре

I. НОРМАЛЬНЫЕ МАТРИЦЫ

Мы будем говорить о квадратных матрицах, элементы которых могут быть комплексными числами (их собственные значения и векторы также могут быть комплексными). Матрицы будем обозначать буквой со шляпкой, например, $\hat{A}$. Нас будет интересовать решение задачи на собственные значения и собственные векторы: $\hat{A}\vec{v}_i = \lambda_i \vec{v}_i$.

Определим несколько понятий, которые нам понадобятся в дальнейшем:

(a) Коммутатор двух матриц: $[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}$. Так как матричное произведение, вообще говоря, зависит от порядка сомножителей, коммутатор в общем случае не равен нулю. Если он равен нулю, говорят, что матрицы коммутируют (т.е. их можно переставлять в произведениях).

(b) Эрмитово сопряжение матрицы — это транспонирование и комплексное сопряжение: $\hat{A}^\dagger \equiv \hat{A}^{T*}$. Порядок этих двух операций не важен, можно написать $\hat{A}^{*T}$.

(c) Для комплексных векторов $\vec{v}_1$ и $\vec{v}_2$ скалярное произведение определяется как $\vec{v}_1^\dagger \vec{v}_2$. Для случая скалярного произведения вектора самого на себя из такого определения получится сумма квадратов модулей компонент.

Мы в данной лекции будем рассматривать так называемые *нормальные матрицы*. Это матрицы, обладающие свойством $[\hat{A}, \hat{A}^\dagger] = 0$. Среди нормальных матриц можно выделить специальные подклассы:

- Эрмитовы матрицы: $\hat{A}^\dagger = \hat{A}$, собственные значения λ_i вещественны. Этот случай особенно важен для физических приложений (например, в квантовой механике).

- Антиэрмитовы матрицы: $\hat{A}^\dagger = -\hat{A}$, собственные значения λ_i чисто мнимые.

- Унитарные матрицы: $\hat{A}^\dagger = \hat{A}^{-1}$, собственные значения λ_i лежат на единичной окружности в комплексной плоскости. Унитарные матрицы можно рассматривать как матрицы поворотов в том смысле, что они не меняют скалярных произведений векторов. Действительно, если на вектора $\vec{v}_1$ и $\vec{v}_2$ подействовать унитарной матрицей $\hat{A}$, то скалярное произведение превратится в $(\hat{A}\vec{v}_1)^\dagger(\hat{A}\vec{v}_2) = \vec{v}_1^\dagger \hat{A}^\dagger \hat{A} \vec{v}_2 = \vec{v}_1^\dagger \vec{v}_2$, т.е. не изменится.

Также можно выделить вещественные варианты тех же классов: симметричные матрицы ($\hat{A}^T = \hat{A}$), антисимметричные матрицы ($\hat{A}^T = -\hat{A}$), ортогональные матрицы ($\hat{A}^T = \hat{A}^{-1}$). Бывают также нормальные матрицы общего вида.

Спектральная теорема утверждает, что матрица $\hat{A}$ является нормальной тогда и только тогда, когда она может быть диагонализована некоторой унитарной матрицей: $\hat{A} = \hat{U}\hat{\Lambda}\hat{U}^\dagger$, где $\hat{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots)$ — диагональная матрица, а $\hat{U}$ — унитарная. При этом оказывается, что λ_i — собственные значения матрицы $\hat{A}$, а столбцы матрицы $\hat{U}$ — нормированные собственные векторы матрицы $\hat{A}$ (расставленные в том же порядке, что и λ_i в $\hat{\Lambda}$).

Важное для нас свойство: собственные векторы нормальной матрицы, соответствующие различным собственным значениям, образуют ортогональный базис (можно нормировать и получить ортонормированный базис). То есть если $\hat{A}\vec{v}_n = \lambda_n \vec{v}_n$ и $\hat{A}\vec{v}_k = \lambda_k \vec{v}_k$, то скалярное произведение $\vec{v}_n^\dagger \vec{v}_k = 0$ при $n \neq k$. Мы будем также использовать обозначения $|v_k\rangle \equiv \vec{v}_k$, $\langle v_n| \equiv \vec{v}_n^\dagger$, $\langle v_n|v_k\rangle \equiv \vec{v}_n^\dagger \vec{v}_k$ (для нас сейчас это просто обозначения, потом дополнительный смысл им придаст квантовая механика).

II. ТЕОРИЯ ВОЗМУЩЕНИЙ

Пусть имеется нормальная матрица $\hat{H}_0$, про которую известны её собственные значения и (ортонормированные) собственные векторы:

$$\hat{H}_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle, \qquad \langle\psi_n^{(0)}|\psi_m^{(0)}\rangle = \delta_{nm}. \tag{2.1}$$

(Мы здесь используем обозначения, аналогичные тем, что будут потом использоваться в квантовой механике.)

Нас интересуют собственные значения E_n и собственные векторы $|\psi_n\rangle$ возмущённой матрицы $\hat{H} = \hat{H}_0 + \hat{\mathcal{V}}$:

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle. \tag{2.2}$$

Можно искать ответы в виде разложений по степеням возмущения:

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \ldots, \tag{2.3}$$
$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + |\psi_n^{(1)}\rangle + |\psi_n^{(2)}\rangle + \ldots \tag{2.4}$$

(Иногда возмущение пишут в виде $\hat{\mathcal{V}} = \epsilon\hat{\mathcal{V}}_0$ с малым параметром ϵ; тогда разложение идет по степеням ϵ.)

На самом деле, вместо прямолинейной записи (2.4) удобнее искать собственные векторы в виде

$$|\psi_n\rangle = \sum_m c_m|\psi_m^{(0)}\rangle, \qquad \text{где } c_m = c_m^{(0)} + c_m^{(1)} + c_m^{(2)} + \ldots, \tag{2.5}$$

— мы будем делать именно так.

Подставляем все в уравнение (2.2):

$$(\hat{H}_0 + \hat{\mathcal{V}})\sum_m c_m|\psi_m^{(0)}\rangle = E_n\sum_m c_m|\psi_m^{(0)}\rangle. \tag{2.6}$$

Учитывая действие невозмущенной матрицы, получаем

$$\sum_m c_m(E_m^{(0)} + \hat{\mathcal{V}})|\psi_m^{(0)}\rangle = \sum_m c_m E_n|\psi_m^{(0)}\rangle. \tag{2.7}$$

Берем проекцию на $\langle\psi_k^{(0)}|$, получаем

$$\boxed{(E_n - E_k^{(0)})c_k = \sum_m V_{km}c_m}\,, \tag{2.8}$$

где матричные элементы V_{km} определены следующим образом:

$$V_{km} \equiv \langle\psi_k^{(0)}|\hat{\mathcal{V}}|\psi_m^{(0)}\rangle. \tag{2.9}$$

Смысл этих матричных элементов необходимо пояснить. Подобно тому как пространственный вектор имеет разные компоненты в разных базисах, матрицу тоже можно рассматривать в разных базисах. Если мы взяли матрицу $\hat{A}$, состоящую из элементов A_{ij}, то можно сказать, что это элементы матрицы в базисе, заданном векторами $\vec{a}_1 = (1,0,0,\ldots)^T$, $\vec{a}_2 = (0,1,0,\ldots)^T$ и т.д. Это означает, что $A_{ij} = \vec{a}_i^\dagger\hat{A}\vec{a}_j$. Но мы можем выбрать другой базис, задаваемый некоторыми векторами $\vec{b}_i$, тогда по отношению к нему матрица будет иметь матричные элементы $A'_{ij} = \vec{b}_i^\dagger\hat{A}\vec{b}_j$. В этом смысле V_{km} — матричные элементы матрицы $\hat{\mathcal{V}}$, вычисленные в базисе векторов $|\psi_k^{(0)}\rangle$.

Далее мы будем рассматривать уравнение (2.8) в разных порядках по возмущению. Детали рассмотрения зависят от того, является ли невозмущённое собственное значение $E_n^{(0)}$ невырожденным (т.е. ему соответствует только один собственный вектор) или вырожденным (т.е. ему соответствуют несколько линейно независимых собственных векторов). Рассмотрим эти ситуации по-отдельности.

III. НЕВЫРОЖДЕННОЕ СОБСТВЕННОЕ ЗНАЧЕНИЕ

Пусть нас интересует собственное значение и собственный вектор с номером n. Тогда из записи (2.5) очевидно, что $c_k^{(0)} = \delta_{kn}$. В нулевом порядке уравнение (2.8) (мы должны сохранить лишь нулевой порядок по V в обеих сторонах уравнения) даёт тривиальное равенство

$$(E_n^{(0)} - E_k^{(0)})\delta_{kn} = 0. \tag{3.1}$$

A. Первый порядок теории возмущений

Первый порядок:

$$E_n = E_n^{(0)} + E_n^{(1)}, \tag{3.2}$$

$$c_k = \delta_{kn} + c_k^{(1)}, \tag{3.3}$$

и уравнение (2.8) в первом порядке даёт

$$E_n^{(1)}\delta_{kn} + (E_n^{(0)} - E_k^{(0)})c_k^{(1)} = \sum_m V_{km}\delta_{mn} = V_{kn}. \tag{3.4}$$

Отсюда при $k = n$ получаем

$$E_n^{(1)} = V_{nn}, \tag{3.5}$$

таким образом мы нашли поправку первого порядка к собственному значению. При $k \neq n$ получаем поправку первого порядка к коэффициентам, определяющим собственный вектор:

$$k \neq n: \qquad c_k^{(1)} = \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}. \tag{3.6}$$

При этом коэффициент $c_n^{(1)}$ остаётся неопределённым. Он должен быть найден из условия нормировки. Записав

$$|\psi_n\rangle = (1 + c_n^{(1)})|\psi_n^{(0)}\rangle + {\sum_m}' \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}}|\psi_m^{(0)}\rangle \tag{3.7}$$

(штрих у знака суммы означает, что суммирование ведется по всем значениям кроме n) и учитывая ортонормированность системы невозмущённых собственных векторов, с точностью до первого порядка получаем

$$\langle\psi_n|\psi_n\rangle = 1 + 2\,\mathrm{Re}\,c_n^{(1)}. \tag{3.8}$$

Отсюда ясно, что мы можем взять

$$c_n^{(1)} = 0. \tag{3.9}$$

Поэтому в результате поправка первого порядка к собственному вектору дается выражением

$$|\psi_n^{(1)}\rangle = {\sum_m}' \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}}|\psi_m^{(0)}\rangle. \tag{3.10}$$

Из формулы (3.10) следует условие применимости теории возмущений. Поскольку поправка к собственному вектору должна быть мала, все коэффициенты в сумме должны быть много меньше единицы:

$$|V_{mn}| \ll |E_n^{(0)} - E_m^{(0)}|. \tag{3.11}$$

B. Второй порядок теории возмущений

Теперь

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)}, \tag{3.12}$$

$$c_k = \delta_{kn} + c_k^{(1)} + c_k^{(2)}, \tag{3.13}$$

и уравнение (2.8) во втором порядке даёт:

$$E_n^{(2)}\delta_{kn} + E_n^{(1)}c_k^{(1)} + (E_n^{(0)} - E_k^{(0)})c_k^{(2)} = \sum_m V_{km}c_m^{(1)}. \tag{3.14}$$

При $k = n$ получаем поправку второго порядка к собственному значению:

$$E_n^{(2)} = \sum_m{}' \frac{V_{nm} V_{mn}}{E_n^{(0)} - E_m^{(0)}}. \tag{3.15}$$

В наиболее физически важном случае эрмитовых матриц будет $V_{mn} = V_{nm}^*$, поэтому в правой части в числителе дроби будет $|V_{nm}|^2$. Это, в частности, означает, что для эрмитовых матриц поправка второго порядка к наименьшему собственному значению всегда отрицательна (т.к. все знаменатели в сумме отрицательны).

При $k \neq n$ из уравнения (3.14) находим

$$c_k^{(2)} = \sum_m{}' \frac{V_{km} V_{mn}}{(E_n^{(0)} - E_k^{(0)})(E_n^{(0)} - E_m^{(0)})} - \frac{V_{kn} V_{nn}}{(E_n^{(0)} - E_k^{(0)})^2}. \tag{3.16}$$

Аналогично тому что было в первом порядке, коэффициент $c_n^{(2)}$ остаётся неопределённым и должен быть найден из условия нормировки. Записывая

$$|\psi_n\rangle = (1 + c_n^{(2)})|\psi_n^{(0)}\rangle + \sum_m{}' c_m^{(1)} |\psi_m^{(0)}\rangle + \sum_m{}' c_m^{(2)} |\psi_m^{(0)}\rangle, \tag{3.17}$$

с точностью до второго порядка получаем

$$\langle \psi_n | \psi_n \rangle = 1 + 2 \operatorname{Re} c_n^{(2)} + \sum_m{}' |c_m^{(1)}|^2. \tag{3.18}$$

Отсюда ясно, что для выполнения условия нормировки мы можем взять

$$c_n^{(2)} = -\frac{1}{2} \sum_m{}' |c_m^{(1)}|^2 = -\frac{1}{2} \sum_m{}' \frac{|V_{mn}|^2}{(E_n^{(0)} - E_m^{(0)})^2}. \tag{3.19}$$

IV. ВЫРОЖДЕННОЕ СОБСТВЕННОЕ ЗНАЧЕНИЕ

Пусть теперь интересующее нас невозмущённое собственное значение $E_n^{(0)}$ вырождено с кратностью s. Это означает, что существует s линейно независимых собственных векторов $|\psi_n^{(0)}\rangle$, $|\psi_{n'}^{(0)}\rangle$, $|\psi_{n''}^{(0)}\rangle$, $\ldots$, соответствующих одному и тому собственному значению $E_n^{(0)}$ невозмущённой матрицы.

Выбор собственных векторов «внутри» вырожденного собственного значения неоднозначен, можно взять любые их линейные комбинации, и при этом по-прежнему будут получаться вектора, отвечающие тому же $E_n^{(0)}$. Но есть так называемые «правильные» векторы нулевого приближения — это такие комбинации $c_n^{(0)}|\psi_n^{(0)}\rangle + c_{n'}^{(0)}|\psi_{n'}^{(0)}\rangle + \ldots$, которые *слабо* меняются под действием возмущения. Иными словами, поправки к таким комбинациям за счёт возмущения должны быть малы в меру малости возмущения. Это нетривиальное требование, для произвольных комбинаций такого не будет.

Можно пояснить эту ситуации на физическом примере. Пусть у нас есть совершенно изотропный ферромагнетик с некоторой намагниченностью $\mathbf{M}$. Мы хотим приложить слабое магнитное поле. Ясно, что величина намагниченности тоже изменится слабо. Но давайте подумаем про намагниченность в полном векторном смысле. Изначально она может быть направлена куда угодно. Когда же мы приложим поле по оси z, намагниченность сначала повернётся и станет направлена по оси z (теперь ей не всё равно, куда смотреть), а затем немного изменится по модулю. То есть изменение вектора $\mathbf{M}$ вообще говоря не мало! Есть поправка нулевого порядка, связанная с перенаправлением $\mathbf{M}$ по z. А если мы, имея в виду дальнейшее включение поля по z, заранее направим $\mathbf{M}$ именно туда, тогда поправка к вектору будет мала. Таким образом, намагниченность, направленная по z — это правильная намагниченность нулевого приближения. С собственными векторами ситуация аналогична. Коэффициенты $c_n^{(0)}$, $c_{n'}^{(0)}$, $\ldots$ никакой малости не содержат — просто, зная какое возмущение затем будет приложено, мы должны заранее повернуть собственные векторы в нужную сторону.

Технически это делается довольно просто. Запишем уравнение (2.8) при $k = n, n', \ldots$ в первом порядке теории возмущений. Поскольку все эти индексы соответствуют одному вырожденному уровню, разность энергий в левой части не содержит нулевого порядка, и мы оставляем там первый порядок $E^{(1)}$. В то же время, в правой стороне первый порядок уже обеспечен за счёт V_{km}, поэтому c_k с обеих сторон нужно брать в нулевом порядке:

$$c_n = c_n^{(0)}, \quad c_{n'} = c_{n'}^{(0)}, \quad \ldots \tag{4.1}$$

$$c_m = 0 \quad \text{при } m \neq n, n', \ldots \tag{4.2}$$

В результате уравнение (2.8) даёт s уравнений

$$E^{(1)} c_n^{(0)} = \sum_{n'} V_{nn'} c_{n'}^{(0)} \tag{4.3}$$

для всех n внутри вырожденного собственного значения. Перепишем это в виде

$$\sum_{n'} (V_{nn'} - E^{(1)} \delta_{nn'}) c_{n'}^{(0)} = 0, \tag{4.4}$$

откуда ясно, что для существования нетривиального решения для коэффициентов $c_n^{(0)}$ необходимо

$$\det[V_{nn'} - E^{(1)} \delta_{nn'}] = 0. \tag{4.5}$$

Это уравнение называется секулярным. Оно полиномиальное, s-ой степени по $E^{(1)}$, поэтому имеет s корней (для эрмитовых матриц можно дополнительно отметить, что все корни вещественны). Сумма его корней равна сумме диагональных элементов $V_{nn} + V_{n'n'} + \dots$ Это можно доказать следующим образом. Обозначим корни λ_i, тогда

$$\begin{aligned} \det[V_{nn'} - E^{(1)} \delta_{nn'}] &= \left(\lambda_1 - E^{(1)}\right) \left(\lambda_2 - E^{(1)}\right) \dots \left(\lambda_s - E^{(1)}\right) = \\ &= \left(-E^{(1)}\right)^s + (\lambda_1 + \lambda_2 + \dots + \lambda_s) \left(-E^{(1)}\right)^{s-1} + \dots \end{aligned} \tag{4.6}$$

С другой стороны, при вычислении детерминанта напрямую получится длинная сумма, но степени $\left(E^{(1)}\right)^s$ и $\left(E^{(1)}\right)^{s-1}$ возникнут только из одного слагаемого этой суммы, содержащего произведение диагональных элементов матрицы, т.е. из слагаемого $(V_{11} - E^{(1)})(V_{22} - E^{(1)}) \dots (V_{ss} - E^{(1)})$. Поэтому

$$\det[V_{nn'} - E^{(1)} \delta_{nn'}] = \left(-E^{(1)}\right)^s + (V_{11} + V_{22} + \dots + V_{ss}) \left(-E^{(1)}\right)^{s-1} + \dots \tag{4.7}$$

Сравнение записей (4.6) и (4.7) доказывает сделанное выше утверждение.

Итак, рассматривая вопрос о коэффициентах, задающих правильные собственные векторы нулевого приближения, мы получили секулярное уравнение (4.5), из которого находятся поправки первого порядка к вырожденному уровню. В результате, вообще говоря, имеет место так называемое снятие вырождения, которое может быть полным (если все корни уравнения различны) либо частичным (вырожденное собственное значение расщепляется, но по крайней мере некоторые расщеплённые собственные значения остаются вырожденными с меньшей кратностью).

Сами же коэффициенты $c_n^{(0)}$, задающие правильные собственные векторы нулевого приближения, нужно находить из уравнения (4.4), подставляя туда по очереди найденные значения поправок $E^{(1)}$.