**Нечто вроде ТЗ для программы оценивания констант равновесия по экспериментальным данным (спектрофотометрия)**

Принципиальная блок-схема программы для расчета констант равновесия имеет следующий вид:

4. Расчет вел-ны минимизационной функции

1. Ввод данных

3. Расчет молярных коэфф. светопоглощения

2. Расчет равновесных концентраций

6. Изменение lg Ki

5. Fmin найден?

Да

7. Расчет погрешностей параметров

8. Вывод данных

1. **Ввод данных**

Входными данными являются:

а) стехиометрическая матрица *νij* размерностью *MxN*, где *M* – число частиц базиса, *N* – число реакций с их участием (число продуктов реакции); *i* – номер реакции, *j* – номер базисной частицы

б) вектор *lg Ki* размерностью *N*;

в) матрица общих концентраций частиц базиса *C(Bjk)* размерностью *MxK*, где *K* – число исследуемых растворов, *k* – номер раствора.

Несложно видеть, что здесь требуется ввод тех же самых данных, что и в предыдущей задаче. Различия начинаются со следующего пункта.

г) матрица оптических плотностей (приведенных к односантиметровой кювете) растворов *Alk* размерностью *LxK*, где *L* – количество длин вол, при которых мы фиксируем оптические плотности, *l* – номер длины волны;

д) матрица абсолютных погрешностей измерения оптических плотностей растворов *δAlk* размерностью *LxK*;

e) матрица молярных коэффициентов светопоглощения εls размерностью *LxS*, где *S* – число частиц (базисных частиц либо продуктов реакции!) для которых известны величины *ε*. Все другие величины *ε* будут в дальнейшем рассчитываться программой.

ж) важные скаляры! – начальный шаг варьирования h1 для расчета по алгоритму Хука-Дживса, а также точность расчета ξ и предельное количество итераций P. У КЕВ параметр уменьшения шага не указан, видимо, нам придется принять его самостоятельно (вроде как он, по умолчанию равен 2).

Кроме того, пользователь должен отметить:

1. Какие константы *lg Ki* необходимо оптимизировать (остальные при этом в дальнейших расчетах варьироваться не будут!);
2. Какие молярные коэффициенты светопоглощения *εlt* необходимо после расчета вывести в качестве выходных данных.
3. **Расчет равновесных концентраций**

Это мы уже умеем делать. Здесь работает та программа, которую ты уже создал.

1. **Расчет молярных коэффициентов светопоглощения**

Молярные коэффициенты светопоглощения связаны с величинами оптической плотности *Alk*, данными в п. 1г и равновесными концентрациями базисных частиц и продуктов реакции следующим уравнением (которое должно применяться к каждому *k*-тому раствору при каждой *l*-ной длине волны!):

 (1),

Чтобы было легче понять, что это уравнение обозначает, давай представим ситуацию, когда у нас в растворе протекает реакция A + B = D (здесь буквы – это обозначения химических веществ). Предположим, что мы регистрировали оптические плотности при 3 длинах волн и приготовили всего 5 растворов с разными общими концентрациями А и В. Предположим также, что ни для одной из частиц (базисной – А и В или продукта реакции D) нам неизвестны коэффициенты молярного светопоглощения.

Берем первый раствор и смотрим на его оптическую плотность при первой выбранной нами длине волны. В этом случае, *А11* = *a1\*ε1a*+*b1\*ε1b*+*d1\*ε1d* (маленькие латинские буквы в данном случае означают равновесные концентрации веществ A, B и D).

А теперь возьмем для примера 4-ый раствор при 2-ой длине волны. Тогда *А24* = *a4\*ε2a*+*b4\*ε2b*+*d4\*ε2d*.

Здесь самое главное не запутаться в номерах, потому что их здесь фигурирует четыре штуки: *j* – номер БЧ, *i* – номер продукта, *l* – номер длины волны и, наконец, *k* – номер раствора.

Решением системы *LхK* уравнений вида (1) (в примере выше, как нетрудно убедиться, получится ровно 15 уравнений) является *L* векторов *εp*, где *p* меняется от 1 до *M+N*. Или же *M+N* векторов *εl*. Или же, другими словами, матрица *ε* размерностью *Lx(M+N)*. В примере выше, как, опять же, несложно убедиться, решениями являются три (по числу длин волн) вектора величин *εр*, где *р* изменяется от 1 до 3. Или же три (по общему числу частиц) вектора величин *εl*, где *l* изменяется от 1 до 3.

Возможно, все же проще будет *L* раз решать систему *K* уравнений с *M+N-S* переменными (где *S* – число частиц, для которых нам известны коэффициенты светопоглощения). Если какие-либо коэффициенты молярного светопоглощения отмечены пользователем как известные в п. 1е, то у нас просто на *S* уменьшается число неизвестных, которые мы должны найти. Вероятно, как в случае предыдущей задачи с равновесными концентрациями, будет проще сразу перенести эти известные слагаемые в левую часть уравнения (1).

Решать системы уравнений вида (1) следует методом наименьших квадратов.

Для этого метода КЕВ дает общее уравнение линейного МНК с учетом неравноточности измерений:

*X* = (*ӐWA*)-1·*ӐWY* (2),

где *X* – вектор искомых параметров (в нашем случае, это вектор εp), *А* – матрица коэффициентов условных уравнений (насколько я понимаю, здесь имеются в виду именно коэффициенты перед неизвестными в уравнении типа (1) – но ни в коем случае не стехиометрическая матрица!), *Y* – вектор экспериментальных величин (в нашем случае при фиксированном значении *l*, это *Ak*), Ӑ означает, что матрицу А транспонировали, *W* – весовая диагональная матрица, элементы главной диагонали которой определяются как:

 (3),

где *s2k* – дисперсия *Ak*.

Поскольку величины *δAlk*, введенные пользователем в п. 1д, чаще всего представляют собой стандартные отклонения величины *Alk*, то *s2k* = (*δAlk*)2.

Разница между формулой (2) и уравнением для решения системы линейных уравнений МНК как раз и заключается во введении диагональной матрицы W.

1. **Расчет величины минимизационной функции**

Минимизируемая функция имеет вид:

 (4),

где экспериментальные значения оптических плотностей вводит пользователь в п. 1г, а расчетные величины оптических плотностей получаются по уравнению (1) при подстановке в него корней (величин ε), найденных МНК в п. 3. Фактически, минимизационная функция представляет собой сумму квадратов невязок в системе линейных уравнений (1). Величины *wl*представляют собой весовые множители – обратные дисперсии величин А (см. п.3, уравнение (3)). КЕВ, впрочем, в другом месте указывает, что в случае равноточных измерений данные величины могут быть приравнены к 1 (видимо, оттого, что в этом случае мы умножаем квадраты разности на одно и то же число, что не имеет особого смысла).

1. **Проверка минимума Fmin**

Фактически, этот и следующий блок представляют собой алгоритм Хука-Дживса для поиска экстремума минимизационной функции. КЕВ не описывает работу данного алгоритма. Поэтому я дам описание АХД, опираясь на данные из других источников (в частности, на оригинальную статью Х-Д, прилагается).

Итак, фактически нашей задачей является поиск минимума функции F(*lg K1*; *lg K2*;…;*lg KN*) (я буду в дальнейшем для краткости обозначать F{lg *Ki*})

5.1. Шаг номер 1. Вычисление Fmin в начальной базовой точке с координатами {*lg Ki*}, заданными пользователем в п. 1б, из которых *u* констант под номерами *q*; …; *t*; …; *z* являются неизвестными и подлежат оптимизации. Это сделано в п. 4.

5.2. Шаг номер 2. Исследующий поиск.

Каждая переменная по очереди изменяется прибавлением шага варьирования. Таким образом, мы фиксируем значения всех *lg K*, кроме логарифма оптимизируемой константы с наименьшим порядковым номером *q* и вычисляем значение функции F{*lg Ki + heq*}, где *eq* - единичный вектор в направлении оси *lg Kq*.

Выполняем проверку, соблюдается ли условие F{*lg Ki + heq*}<F{*lg Ki*}.

Если да, то мы заменяем старую базовую точку с координатами {*lg Ki*} новой с координатами {*lg Ki + heq*}. После этого мы фиксируем все значения *lg K* в этой новой точке за исключением переменной со следующим по порядку номером *t*, для которой проделываем ту же операцию.

Если нет, то мы вычисляем значение функции F{*lg Ki – heq*}.

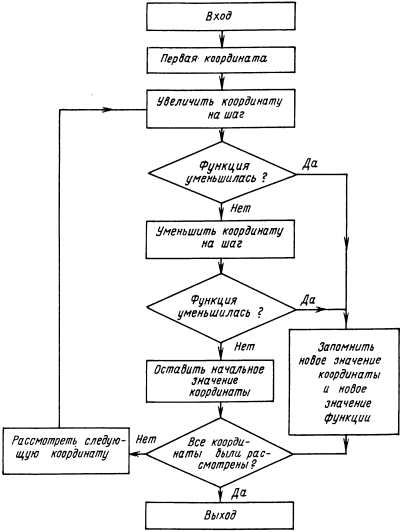
Выполняем проверку, соблюдается ли условие F{*lg Ki – heq*}<F{*lg Ki*}.

Если да, то мы заменяем старую базовую точку с координатами {*lg Ki*} новой с координатами {*lg Ki – heq*}. После этого мы фиксируем все значения *lg K* в этой новой точке за исключением переменной со следующим по порядку номером *t*, для которой проделываем ту же операцию.

Если поиск по обоим направления шага по переменной lg Kq не привел к выполнению условия F{*lg Ki + heq*} или F{*lg Ki – heq*}<F{*lg Ki*}, то мы оставляем переменную с наименьшим порядковым номером *lg Kq* в покое и переходим к следующей по порядку переменной, для которой выполняем исследующий поиск таким же образом.

После того, как мы выполним исследующий поиск по всем *u* переменным, у нас будет новая базовая точка с координатами {*lg Ki’*}.

Блок-схема исследующего поиска выглядит следующим образом:



5.3. Проверяем условие {*lg Ki*}={*lg Ki’*}.

Если оно соблюдается, т.е. мы не добились уменьшения F, необходимо повторить исследующий поиск вокруг точки с координатами {*lg Ki*} (т.е. начальной базовой точки), но с уменьшенной длиной шага варьирования h’=h/α, где α – параметр сжатия (можно принять его равным 2).

Если условие {*lg Ki*}={*lg Ki’*} не соблюдается, необходимо перейти к поиску по образцу.

5.4. При поиске по образцу используется информация, полученная в процессе исследования, и минимизация функции завершается поиском в направлении, заданном образцом. Эта процедура производится следующим образом:

Разумно двигаться в направлении от начальной базовой точке к базовой точке номер 1, поскольку в этом направлении удалось добиться уменьшения F. Координаты этой «промежуточной» точки вычисляются по формуле:

{*lg Ki’’*} = {*lg Ki*} + 2({*lg Ki’*}-{*lg Ki*}) (5);

Или, в общем случае:

{*lg Kij*} = {*lg Kij-2*} + 2({*lg Kij-1*}-{*lg Kij-2*}) (6);

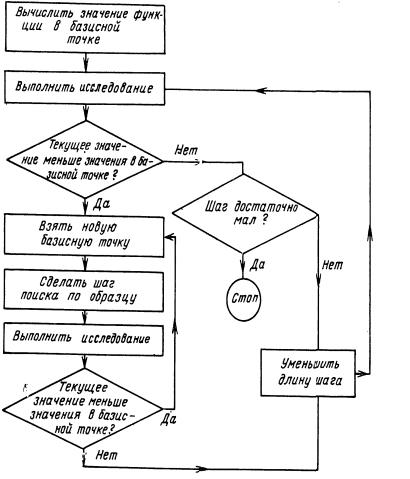
После этого, необходимо выполнить исследующий поиск вокруг этой полученной промежуточной точки. Единственное, в чем будет отличие – шаг варьирования на этом этапе уменьшать не нужно!

Если в результате исследующего поиска вокруг промежуточной точки удалось получить новую точку {*lg Ki’’’*} для которой значение F<F{*lg Ki’*}, то проводим поиск по образцу от точки {*lg Ki’*} к точке {*lg Ki’’’*}.

Если условие F{*lg Ki’’’*}<F{*lg Ki’*} не выполняется, то необходимо перейти к исследующему поиску вокруг точки {*lg Ki’*}.

Все это безобразие повторяется до тех пор, пока величина h не будет уменьшена до наперед заданной точности ξ (см. п. 1ж).

Блок-схема АХД в общем виде выглядит следующим образом:



Полезные линки к данному разделу:

А) <https://www.intuit.ru/studies/courses/1020/188/lecture/4931>

Б) <https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%A5%D1%83%D0%BA%D0%B0_%E2%80%94_%D0%94%D0%B6%D0%B8%D0%B2%D1%81%D0%B0>

В) <https://math.semestr.ru/optim/example-hook-jeeves.php>

Г) Оригинальная статья Хука и Дживса (прикрепил отдельно)

Обрати внимание, что в нашем случае Fmin не является функцией от {*lg Ki*} в явном виде. Для того, чтобы находить значения Fmin, нам всякий раз нужно выполнять блоки 2 и 3 (расчет равновесного состава и молярных коэффициентов светопоглощения).

Я полагаю, что АХД уже реализован в виде отдельной библиотеки. Возможно, что даже КЕВ уже его откуда-то взял, а не прописывал самостоятельно, поскольку описания алгоритма у него нет.

**7. Расчет погрешностей оптимизированных параметров**

В блоке оценки погрешностей рассчитывает ковариационная матрица случайных ошибок:

 (7),

где Fmin – найденное значение минимума минимизационной функции, М – число степеней свободы (будем считать, что наши измерения равноточны, тогда М определяется как число независимых и одинаково распределенных измерений минус число независимых переменных, т.е. в нашем случае, это будет *K*-*u*, где *K* – число растворов, *u* – число оптимизируемых констант. Важно: именно *K*, но не *KxL*, ибо вылет по одной длине волны почти наверняка означает, что исследователь ошибся с раствором в целом, иными словами, измерения в одном растворе при разных длинах волн не являются независимыми.), U – матрица Якоби, в которой будет *K* строк (по числу независимых измерений, сиречь, растворов) и *u* столбцов (по числу оптимизируемых констант) вида:

| *dA1/d lg Kq* … *dA1/d lg Kz* |

| … … … | (8),

| *dAK/d lg Kq* … *dAK/d lg Kz*|

Частные производные – элементы Якобиана – рассчитываются конечно-разностным методом согласно (9):

 (9),

В уравнении (9) вычисляются значения Ak при значениях оптимизированных констант плюс и минус заданная наперед погрешность расчета констант (см. п. 1ж – важные скаляры).

В этом же блоке необходимо провести проверку чувствительности минимизационной функции F к параметрам *lg Ku*. Если выполняется условие:

(10),

то необходимо вывести на экран оповещение о нечувствительности минимизационной функции к параметру *lg Ku*. Кроме того, если F’ или F’’ меньше, чем F(lg K1; …; lg Ku; …; lg KN), то необходимо вывести на экран сообщение, что минимум F не найден.

Почему такое бывает, написано в примечаниях в конце документа.

Диагональные элементы ковариационной матрицы (7) будут представлять собой дисперсии оптимизируемых параметров. Таким образом, стандартные отклонения искомых констант lg Ku будут рассчитываться как:

 (11),

В случае одновременной оптимизации более, чем одной константы равновесия, недиагональные элементы ковариационной матрицы:

 (12),

дают возможность рассчитывать коэффициенты парной корреляции искомых параметров lg Ku:

 (13),

Если все было выполнено правильно, то коэффициенты корреляции должны принимать значения из отрезка [-1;1]. Смысл коэффициентов корреляции в том, что если он близок к 1 по абсолютной величине, две данных оптимизируемых константы, фактически, не были определены независимо друг от друга. Могут быть получены только сведения об их линейной комбинации. При неверном выборе экспериментального плана константы сильно коррелированы для случая одновременного расчет констант протонирования и комплексообразования лиганда, ступенчатые константы комплексообразования, реакции конкурирующего комплексообразования и т.д. Таким образом, информация о коэффициентах корреляции является важной для исследователя и должна выводиться на экран.

**8.** **Вывод данных**

После завершения расчета, на экран должны выводиться следующие значения:

А) Рассчитанные при помощи уже оптимизированных констант равновесия величины равновесных концентраций;

Б) Рассчитанные при помощи уже оптимизированных констант равновесия значения оптических плотностей при всех длинах волн и их отклонения от экспериментальных значений (абсолютное и относительное);

В) Сумма квадратов отклонений и значение Fmin;

Г) Оптимизированные значения lg Ku вместе со стандартными отклонениями. Например, lg K(7) = 10.48±0.52

Д) Если были оптимизированы несколько констант, должны выводиться коэффициенты парной корреляции.

Е) Рассчитанные при помощи уже оптимизированных констант равновесия значения молярных коэффициентов светопоглощения. В том случае, если какие-либо коэффициенты ε получились отрицательными (что не имеет физического смысла), на экран должно выводиться предупреждение, что такой-то коэффициент отрицательный.

Ж) Если минимизационная функция не чувствительна к к.-либо константе или же минимум не найден (вообще или за отпущенное число итераций), должны выдаваться соответствующие предупреждения (см. п. 7).

**ПРИМЕЧАНИЯ К П.7**

Прямая задача химических равновесий (ПЗХР) – расчет равновесного состава смеси – которую мы решили до того, всегда имеет решение. В отличие от ПЗХР, обратная задача – оценивание констант равновесия по экспериментальным данным может не иметь решения (и весьма часто так и делает).

Можно выделить следующие группы причин, по которым не удается оптимизировать константы равновесия.

I. Особенности расчета по АХД:

А) неудачное начальное приближение искомых констант. Если начальное приближение находится в локальном минимуме, дно которого находится «глубже», чем заданный шаг варьирования, АХД не позволит выбраться из этого локального минимума, поскольку в нем предусмотрено сокращение шага варьирования. Решение проблемы: выполнение расчета из разных начальных приближений до получения сходящихся результатов.

Б) задание в расчет большого количества констант равновесий. Фактически, АХД в случае оптимизации N параметров сводится к задаче поиска глобального минимума функции в N+1-мерном пространстве, соответственно, чем больше N, тем труднее задача, и тем сложнее доказать исследователю, что он действительно обнаружил глобальный экстремум. Решение проблемы: не рекомендуется оптимизировать более 2-3 констант одновременно. В идеале, следует оптимизировать 1 константу за раз. Если неизвестных параметров слишком много, следует определить некоторые из них независимо в отдельных экспериментах.

В) разным по размеру наборам искомых параметров соответствуют одинаковые значения минимизационной функции. Решение проблемы: следует стремиться к минимизации стехиометрической матрицы в соответствии с принципом бритвы Оккама. Кроме того, следует избегать введения в стехиометрическую матрицу процессов, единственной задачей которых является уменьшение Fmin, поскольку уменьшение дисперсии адекватности математической модели не дает ни малейших оснований полагать ее истинной.

Г) разным по существу, но одинаковым по размеру наборам искомых параметров соответствуют одинаковые значения минимизационной функции. Решение проблемы: дискриминация некоторых из имеющихся математических моделей с опорой на дополнительные эксперименты, общие физико-химические знания и элементарный здравый смысл.

II. Экспериментальные:

А) наличие грубых промахов в экспериментальных данных, задаваемых в расчет. Решение проблемы: повторение эксперимента и анализ грубых промахов при помощи инструментов математической статистики (напр., по критерию Диксона).

Б) неудачный выбор общих концентраций. Если выход продукта реакции с искомой константой слишком мал или, напротив, слишком велик, это приводит, в первом случае, к нечувствительности функции к искомому параметру, а во втором случае – к невозможности нахождения минимума. Практика показывает, что выход продукта должен составлять не менее 15% и не более 85% для успешного расчет константы равновесия. Решение проблемы: исследователь с опорой на дополнительные эксперименты и литературные данные должен примерно представлять порядок искомых величин.

В) отказ от задания молярных коэффициентов светопоглощения в расчетную схему. Решение проблемы: следует определять независимо величины ε в дополнительных экспериментах (напр., методом градуировочного графика). Чем более подробную информацию исследователь задаст в расчет, тем выше шансы на успешный и небессмысленный результат расчета.

Г) значительная неопределенность других констант равновесия, задаваемых как известные. Решение проблемы: тщательный анализ литературных данных с целью выбора наиболее надежных и достоверных данных либо изменение плана эксперимента с целью исключения процесса с неопределенной константой.

Д) работа в диапазоне длин волн или оптических плотностей, где результат существенно искажен вследствие особенностей оптической системы спектрофотометра или применяемого растворителя. Решение проблемы: нежелательно выходить за рамки оптимальных значений А, а также учитывать поглощения растворителя (например, ДМСО делает неинформативной область λ<240 нм).