ОПТИМИЗАЦИЯ НЕЛИНЕЙНЫХ СИСТЕМ

С.А. Родионов, Е.И. Гутман

1.1. ПОСТАНОВКА ЗАДАЧИ

При решении различного рода инженерных задач, связанных с проектированием приборов, механизмов, технологических и производственных процессов, почти неизбежно возникает необходимость в оптимизации нелинейных систем, т. е. в нахождении таких значений свободных конструктивных параметров, которые обеспечивали бы оптимальную балансировку характеристик качества проектируемого прибора или механизма. Так, например, при синтезе пространственных механизмов необходимо найти такие значения параметров механизма x_1, x_2, x_3, которые обеспечивали бы наименьшее отклонение реальной функции, воспроизводимой механизмом, от заданной в некотором множестве точек. При расчете оптических систем задача заключается в определении конструктивных параметров системы, обеспечивающих минимальные значения аберраций.

Ранее, при «ручном» проектировании сравнительно несложных механизмов и приборов, для решения возникающих в процессе проектирования задач оптимизации было достаточно опыта и искусства конструктора. В настоящее время, при широком внедрении вычислительной техники в проектирование, усложнении проектируемых объектов, повышении требований к качеству проекта и скорости его выполнения, кустарный подход к решению задач оптимизации при проектировании недопустим.

Для эффективного использования вычислительной техники, в частности универсальных электронных вычислительных машин (ЭВМ), необходимо знание математических основ оптимизации и наличие в математическом обеспечении ЭВМ универсальных стандартных оптимизирующих программ.

К сожалению, в математическом обеспечении большинства современных ЭВМ такие программы отсутствуют, и конструктор, использующий ЭВМ в своей работе, вынужден разрабатывать программы оптимизации на доступном ему алгоритмическом языке, применительно к своей конкретной задаче.

В настоящей главе авторы, основываясь на опыте разработки универсальной оптимизирующей программы для математического обеспечения ЭВМ «Минск-22» [5], поставили перед собой цель изложить в доступной для конструктора форме основные принципы оптимизации нелинейных систем, чтобы помочь ему правильно оценить имеющиеся оптимизирующие программы и методы, эффективно и сознательно их использовать и, в случае необходимости, составить оптимизирующую программу для своей задачи.

Глава книги "Расчет и конструирование механизмов и деталей приборов". Л. Машиностроение. 1975. с.5-37

В общем виде задачу оптимизации можно сформулировать следующим образом. Пусть имеется система m нелинейных функций y_i от n параметров (переменных) ω_j:

$$\left.\begin{aligned} y_1 &= y_1(\omega_1, \omega_2 \ldots \omega_n); \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= y_m(\omega_1, \omega_2 \ldots \omega_n); \end{aligned}\right\} \qquad (1.1)$$

Если считать, что параметры (ω_i и функции y_i являются действительными, удобно ввести в рассмотрение действительные векторные пространства: n–мерное пространство R_x^n параметров и m–мерное пространство R_f^m функций. Тогда можно сказать, что состояние системы описывается вектором W в пространстве параметров R_x^n и вектором Y в пространстве функций R_f^m.

Функциональная зависимость (1.1) отображает пространство R_x^n в пространстве R_f^m и может быть записана в матричной форме

$$Y = Y(W), \qquad (1.2)$$

где $W = \begin{pmatrix} \omega_1 \\ \omega_2 \\ . \\ . \\ . \\ \omega_n \end{pmatrix}$ – матрица-столбец (вектор) параметров;

$Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_m \end{pmatrix}$ – матрица-столбец (вектор) функций.

Оптимизация заключается в нахождении такой точки W_{opt} в пространстве параметров, в которой некоторая оценка φ функций Y, т.е. функционал на пространстве функций R_f^m был бы оптимален (минимален или максимален, в соответствии с предварительной договоренностью). Для определенности в дальнейшем будем говорить о нахождении минимума φ, т.е. о минимизации нелинейной системы. Функционал φ обычно называют оценочной функцией, целевой функцией, функцией качества. Мы будем придерживаться первого наименования.

В общем случае не все точки пространства параметров соответствуют физически реализуемой или конструктивно приемлемой системе, тогда в пространстве параметров задается область (подпространство) Ω, в которой должно находиться искомое решение W_{opt}. Эта область задается системой r неравенств (ограничений), в общем случае также нелинейных относительно параметров:

$$\left.\begin{array}{l} b_1 = b_1(\omega_1, \omega_2 \ldots \omega_n) > 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_r = b_r(\omega_1, \omega_2 \ldots \omega_n) > 0 \end{array}\right\} \qquad (1.3)$$

или в матричной форме

$$B(W) > 0, \qquad (1.4)$$

где $B = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_r \end{pmatrix}$ – матрица-столбец ограничений;

$0 = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix}$ – нулевая матрица-столбец.

Вопросы контроля ограничений в процессе оптимизации представляют значительный интерес как с теоретической, так и практической точек зрения, однако мы их рассматривать не будем, считая, что оптимизация происходит во всем пространстве R_x^n.

Для наибольшей общности не будем требовать явного задания функций y_i, ограничившись требованиями однозначности, непрерывности и дифференцируемости. Предположим, что имеется некоторый алгоритм, или устройство для получения значений функций y_i, т.е. вектора Y в любой точке W. Процесс получения этих значений, т.е. обращение к такому алгоритму, будем называть «пробой» системы. Содержание и объем пробы определяются конкретной технической задачей оптимизации, т.е. являются внешними по отношению к алгоритму оптимизации.

Предположим также, что имеется алгоритм «проба производных», который в любой точке W пространства параметров дает значения частных

производных функций по параметрам, т.е. значения матрицы производных A:

$$A = \left\{ \frac{\partial y_i}{\partial w_j} \right\}; i = \overline{1,m}; j = \overline{1,n}$$

В простейшем случае, при вычислении производных методом конечных разностей, алгоритм проба производных состоит из n раз повторенного алгоритма проба.

В дальнейшем примем, что указанные пробы являются единственным средством получения информации об оптимизируемой системе. Поскольку содержание и объем пробы даются нам априорно и часто требуют большого количества вычислений, эффективность того или иного метода оптимизации оценивается количеством затраченных в процессе поиска оптимальной точки W_{opt} проб, называемом «ценой поиска».

Итак, сформулируем теперь задачу оптимизации. Имеется некоторый алгоритм проба, обращаясь к которому можно любой точке W n–мерного векторного пространства параметров R_x^n сопоставить точку Y m–мерного пространства функций R_f^m, и алгоритм проба производных (в простейшем случае состоящий из n проб), ставящий в соответствие точке W матрицу производных A. Задан, кроме того, вид некоторого функционала $\varphi = \varphi(Y)$ – оценочной функции, отображающий пространство функций R_f^m на множество неотрицательных действительных чисел A_f. Требуется построить процесс, который, затратив наименьшее количество проб, позволил бы определить в пространстве параметров точку W_{opt}, в которой φ минимален;

$$Y = Y(W); W \in R_x^n; Y \in R_f^m; \varphi(Y) \in \Lambda_f \qquad (1.5)$$

Найти W_{opt} такое, что $\varphi(W_{opt}) \le \varphi(W)$ для любого $W \in R_x^n$.

1.2. ВЫБОР ОЦЕНОЧНОЙ ФУНКЦИИ И НОРМИРОВАНИЕ ПРОСТРАНСТВ ПАРАМЕТРОВ И ФУНКЦИЙ

Прежде чем рассматривать процесс поиска минимума, необходимо обсудить возможный вид функционала φ – оценочной функции. При выборе этого функционала необходимо исходить из следующих требований. Из смысла оптимизации вытекает основное требование наличия у φ хотя бы одного минимума и адекватности φ критерию качества оптимизируемой системы. Другими словами, меньшая величина φ должна соответствовать лучшей системе.

Для удобства анализа и построения алгоритма оптимизации желательно иметь по возможности простую связь φ с функциями Y.

Наиболее естественным и, по-видимому, наиболее удобным является выбор φ в виде квадрата длины вектора Y. Для определения квадрата длины в пространстве R_f^m функций необходимо ввести метрику

$$\varphi = d_f^2 = \sum_{i=1}^{m}\sum_{k=1}^{m}\mu_{ik} y_i y_k \qquad (1.6)$$

где μ_{ik} – элементы матрицы – метрики пространства R_f^m – симметрической, положительно определенной матрицы M_f.

В матричной форме предыдущее выражение можно записать следующим образом:

$$\varphi = d_f^2 = Y^T M_f Y \qquad (1.7)$$

где индекс т означает транспонирование, т. е. замену строк матрицы столбцами, и наоборот. При транспонировании матрица–столбец Y переходит в матрицу-строку $Y^T = (y_1, y_2, ...y_m)$. Если M_f – единичная матрица, то d_f^2 представляет собой просто сумму квадратов отдельных функций

$$\varphi = d_f^2 = Y^T Y = \sum_{i=1}^{m} y_i^2 \qquad (1.8)$$

Однако в силу того, что отдельные функции y_i могут иметь различные физические размерности и различный вклад в качестве оптимизируемой системы, наиболее распространенной является диагональная метрика

$$M = \begin{pmatrix} 1/\Delta y_1^2 & & & 0 \\ & 1/\Delta y_2^2 & & \\ & & . & \\ & & & . \\ 0 & & & 1/\Delta y_m^2 \end{pmatrix} \qquad (1.9)$$

В этом случае φ – сумма квадратов функций, деленных на некоторые масштабы Δy_i, размерность которых совпадает с размерностью соответствующих функций, а величина их обратно пропорциональна «вкладам», «весам» функций в качество системы. Сама оценочная функция φ при этом безразмерна.

Введение такой диагональной метрики можно также рассматривать как переход от ненормированного пространства функций Y к нормированному F путем деления каждой функции y_i, на свой масштаб Δy_i. В таком нормированном пространстве все функции безразмерны, а метрика единична

$$\varphi = \sum_{i=1}^{m}\left(\frac{y_i}{\Delta y_i}\right)^2 = \sum_{i=1}^{m} f_i^2 = F^T F\,. \qquad (1.10)$$

Пространство с любой недиагональной метрикой также может быть преобразовано в нормированное пространство с единичной метрикой. Пусть M_f – произвольная матрица-метрика (симметрическая, положительно определенная). Из матричной алгебры [2, 7] следует, что такую матрицу путем приведения к диагональной форме можно представить в виде произведения

$$M_f = U^T \Lambda U, \tag{1.11}$$

где U – ортонормированная матрица собственных векторов матрицы M_f, а Λ – диагональная матрица положительных собственных значений.

Переход от ненормированного пространства функций Y к нормированному F в этом случае можно представить как умножение вектора Y слева на матрицы U и $\Lambda^{1/2}$

$$F = \Lambda^{\frac{1}{2}} UY \tag{1.12}$$

Легко проверить, что сумма квадратов нормированных функций в этом случае совпадает с (1.7)

$$\begin{aligned} \varphi = \sum_{i=1}^{m} f_i^2 = F^T F = \left(\Lambda^{\frac{1}{2}} UY \right)^T \left(\Lambda^{\frac{1}{2}} UY \right) = Y^T U^T \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} UY \\ = Y^T U^T \Lambda UY = Y^T M_f Y \end{aligned} \tag{1.13}$$

Переход к нормированному пространству в общем случае, в соответствии с выражением (1.12), отличается от случая диагональной метрики (1.9) тем, что, кроме масштабирования (деления каждой функции на масштаб $\Delta y_i = \lambda_i^{1/2}$), описываемого умножением Y на диагональную матрицу $\Lambda^{1/2}$, добавляется еще поворот осей координат в пространстве, описываемый умножением на ортонормированную матрицу U.

Обычно недиагональная метрика не употребляется в практических задачах оптимизации, и наиболее распространенной является диагональная метрика (1.9), т. е. масштабирование функций.

Так как в дальнейшем нам понадобится понятие длины вектора и в пространстве параметров, то, проводя аналогичные рассуждения, введем метрику M_x в пространстве параметров W, определив длину в этом пространстве следующим образом:

$$d_x^2 = W^T M_x W = \sum_{j=1}^{n} \sum_{k=1}^{n} \nu_{jk} \omega_j \omega_k$$

После введения метрики M_x, т.е. масштабирования параметров ω_j в соответствии с заданными масштабами $\Delta\omega_j$ и в общем случае – поворота осей в пространстве параметров, мы можем перейти от неносрмированного

пространства W к нормированному пространству X, в котором метрика единична:

$$d_x^2 = \sum_{j=1}^{n} x_j^2 = X^T X; X = \Lambda_x^{\frac{1}{2}} U_x W . \qquad (1.14)$$

В дальнейшем будем рассматривать процесс оптимизации применительно к нормированным пространствам как параметров, так и функций, в которых метрика единична. В таком случае задача оптимизации заключается в нахождении такой точки $X_{\min}$ нормированного пространства параметров, в которой оценочная функция $\varphi = F^T F$. равная квадрату длины вектора нормированных функций, минимальна.

Будем считать, что нормирование производится отнесением каждой функции y_i и параметра ω_j к своим масштабам Δy_i и $\Delta \omega_j$. выбор которых является внешним по отношению к оптимизатору и предоставляется пользователю (конструктору). Правильный выбор этих масштабов, а также самих функций y_i заключается как раз в обеспечении адекватности оценочной функции φ критерию качества системы.

1.3. КЛАССИФИКАЦИЯ МЕТОДОВ ПОИСКА

В дальнейшем мы увидим, что от выбора метрики (т. е. масштабов) в пространстве параметров существенно зависит эффективность того или иного метода оптимизации.

Существуют более или менее подробные классификации методов поиска минимума нелинейных систем [1, 4, 6]. Для дальнейшего изложения нам необходимо только подчеркнуть самые принципиальные различия.

Глобальные методы поиска отличаются тем, что вся область пространства параметров, в которой ищется минимум, покрывается множеством пробных точек, в соответствии с выбранным законом, и на основании полученной информации строится глобальная модель поведения системы во всей области. Затем при помощи анализа этой модели указывается точка предполагаемого минимума или область, в которой она может находиться, меньшая по сравнению с исходной областью.

Далее процесс повторяется в этой меньшей области для уточнения положения минимума и т. д. Цель глобальных методов заключается в поиске глобального, самого глубокого минимума, что, вообще говоря, крайне желательно для многоэкстремальных систем. Однако применение глобальных методов поиска к многомерным пространствам параметров наталкивается на серьезные трудности, заключающиеся в том, что количество необходимых пробных точек возрастает как показательная функция от размерности n пространства параметров и в том, что с увеличением числа пробных точек усложняется построение и анализ модели системы. Это приводит к тому, что глобальные методы поиска с успехом

могут применяться только для небольших размерностей пространства параметров ($n \le 2$).

Локальные методы поиска представляют собой итеративные процессы, каждый *k*-й шаг которых состоит из следующих этапов:

1) анализ поведения системы в окрестности начальной точки $X_0^{(k)}$ и построение простой математической модели системы;

2) построение одноразмерной траектории $\Delta X(s)$ движения в сторону предполагаемого минимума;

3) выбор точки на траектории $X_l^{(k)} = X_0^{(k)} + \Delta X(s)$ (выбор параметра s – длины шага), которая является конечной точкой данного шага;

4) проверка условия окончания поиска, т. е. проверка, не является ли найденная точка достаточно близкой к минимуму.

Если четвертое условие не выполняется, точка $X_l^{(k)}$ принимается за начальную точку $X_0^{(k+1)}$ следующего шага и процесс повторяется. При этом предполагается, что удовлетворяется условие сходимости $\varphi_{k+1} = \varphi\left(X_l^{(k)}\right) < \varphi\left(X_0^{(k)}\right) = \varphi_k$ т. е. в процессе оптимизации происходит монотонное уменьшение оценочной функции φ и, следовательно, процесс сходится к локальному минимуму $X_{\min}$ (рис. 1.1).

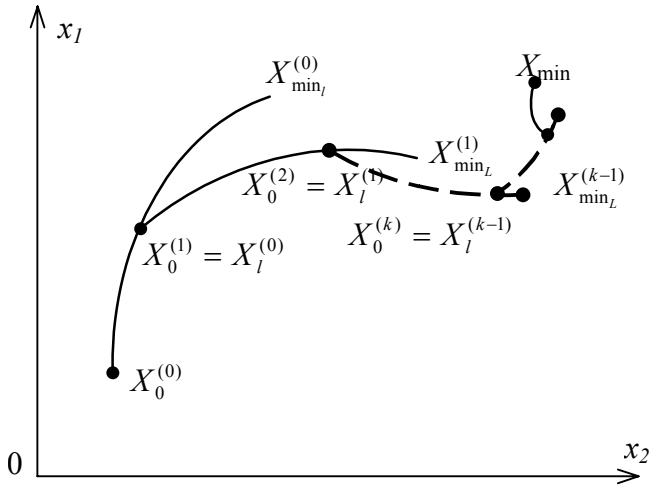

Рис.1.1.

Локальные методы поиска гораздо более экономичны и эффективны для унимодальных (имеющих не более одного минимума в рабочей области) систем, а для сложных систем с большим количеством параметров в настоящее время являются основными рабочими методами.

Проблема поиска глобального минимума многоэкстремальных систем может быть решена многократным применением локального поиска из различных начальных точек, выбираемых по определенному алгоритму, т. е. некоторой комбинацией глобальных и локальных методов поиска [4].

Как локальные, так и глобальные методы могут быть либо случайными, либо детерминированными. Случайные методы отличаются значительной

простотой реализации и для несложных систем весьма эффективны [1]. Однако детерминированные методы, как показывает практика, для сложных систем, у которых проба занимает большой объем вычислений, оказываются более эффективными, сложность метода окупается здесь за счет экономии количества необходимых проб.

В дальнейших параграфах будут рассмотрены основные аспекты построения детерминированных локальных методов поиска.

1.4. АНАЛИЗ ПОВЕДЕНИЯ СИСТЕМЫ В НАЧАЛЬНОЙ ТОЧКЕ И ПОСТРОЕНИЕ ЛОКАЛЬНОЙ МОДЕЛИ СИСТЕМЫ

Если функции $f_i(i=\overline{1,m})$ аналитические в начальной точке X_0, то полную модель поведения системы в окрестности этой точки дает разложение ее в ряд Тейлора. Однако проба и проба производных дают информацию, достаточную для построения только первых двух членов ряда. В принципе можно потребовать информацию о вторых, третьих и т. д. производных, которую всегда можно получить численным методом, проделав необходимое количество проб. Однако количество проб, нужное для этого, пропорционально n^q, где n – размерность пространства параметров; q – порядок производной, т. е. резко возрастает с увеличением q; кроме того, полученная модель получается сложной для использования, требует большого объема памяти; возрастают погрешности производных. Все это приводит к тому, что попытки использования вторых и тем более третьих производных себя не окупают [13]. Таким образом, приходим к линейной модели поведения системы в окрестности исходной точки X_0

$$f_{Li} = f_i(X_0) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_j - x_{0j}) \ \ (i=\overline{1,m}). \tag{1.15}$$

В матричной форме линейная модель запишется в виде

$$F_L = F_0 + A\Delta X, \tag{1.16}$$

где $F_0 = F(X_0)$ – вектор значений функций в исходной точке; $\Delta X = X - X_0$; $A = \left\{ \frac{\partial f_i}{\partial x_j} \right\}$ – матрица частных производных в точке X_0.

Напомним, что здесь мы уже имеем дело с нормированными пространствами параметров и функций.

Рассмотрим выражение для оценочной функции φ_L линейной модели. Пользуясь тем, что $(A\Delta X)^T = \Delta X^T A^T$, получим

$$\varphi_L = F_L^T F_L = (F_0 + A\Delta X)^T (F_0 + A\Delta X) =$$
$$= F_0^T F_0 + F_0^T A\Delta X + \Delta X^T A^T F_0 + \Delta X^T A^T A\Delta X.$$

Заметим, что слагаемые предыдущего выражения есть числа, т.е. матрицы размерности 1×1, которые при транспонировании не изменяются, поэтому можно написать $F_0^T A\Delta X = \Delta X^T A^T F_0$. Тогда

$$\varphi_L = F_0^T F_0 + 2\Delta X^T A^T F_0 + \Delta X^T A^T A\Delta X.$$

Введем обозначения: $N = A^T F_0$ – матрица-столбец размерности $n\times1$; $M = A^T A$ – квадратная, симметрическая, неотрицательно определенная матрица размерности $n\times n$ (порядка n). Размерности этих матриц вытекают из правил матричного умножения, а симметрия матрицы M легко показывается. Действительно, матрица называется симметрической, если она совпадает с транспонированной: $M^T = M$. Для транспонированной матрицы M^T получаем

$$M^T = (A^T A)^T = A^T (A^T)^T = A^T A = M$$

Неотрицательная определенность M при любой матрице A также может быть доказана [7].

Используя обозначения M и N, запишем выражение для линейной модели φ_L.

$$\varphi_L = \varphi_0 + 2\Delta X^T N + \Delta X^T M\Delta X \qquad (1.17)$$

где $\varphi_0 = F_0^T F_0$ – значение оценочной функции в начальной точке.

В заключение данного параграфа рассмотрим градиент оценочной функции. Оценочная функция, как следует из ее определения, есть скалярная функция точки X пространства параметров, т.е. скалярное поле. Поверхности $\varphi = c$ есть поверхности уровня φ (рис. 1.2).

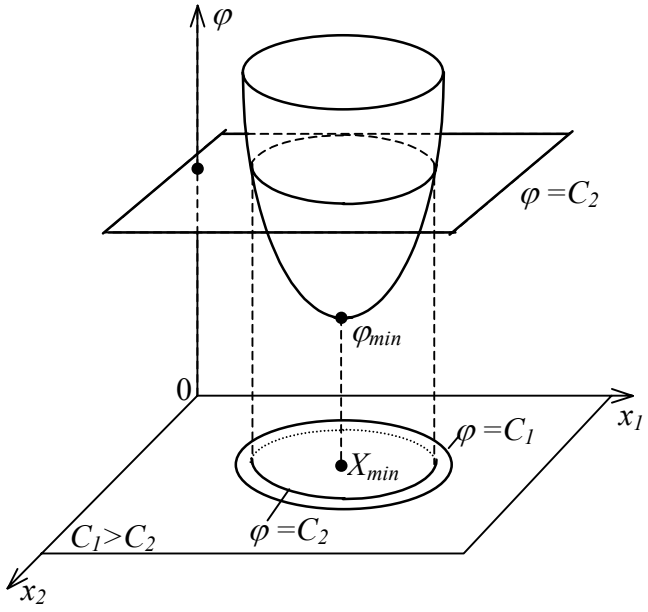

Рис.1.2

Выражение (1.17) показывает, что для оценочной функции линейной модели поверхности уровня есть концентрические эллипсоиды (в n–мерном пространстве параметров – гиперэллипсоиды).

В каждой точке скалярного поля может быть определен вектор, называемый градиентом, который направлен нормально к поверхности уровня, т.е. в сторону быстрейшего возрастания функции φ, а величина его пропорциональна производной оценочной функции по направлению градиента. Из теории скалярного поля следует, что проекции

градиента равны частным производным оценочной функции по соответствующим параметрам:

$$\overrightarrow{grad\varphi} = \left\{ \frac{\partial \varphi}{\partial x_j} \right\} \quad j = \overline{1,n} \tag{1.18}$$

Введем вместо $\overrightarrow{grad\varphi}$ вектор G, равный половине градиента:

$G = \frac{1}{2}\overrightarrow{grad\varphi}$. Впоследствии, из соображений удобства, G будем называть градиентом. Дифференцируя выражение (1.13) для оценочной функции, получим

$$\frac{\partial \varphi}{\partial x_j} = \frac{\partial}{\partial x_j}(F^T F) = F^T \left[\frac{\partial}{\partial x_j}(F) \right] + \left[\frac{\partial}{\partial x_j}(F^T) \right] F = \left\{ F^T \left[\frac{\partial}{\partial x_j}(F) \right] \right\}^T +$$

$$+ \left[\frac{\partial}{\partial x_j}(F) \right]^T F = 2 \left[\frac{\partial}{\partial x_j}(F) \right]^T F$$

(оба слагаемых есть числа, т. е. матрицы 1×1, поэтому при транспонировании не изменяются),

Заметим, что $\frac{\partial}{\partial x_j}(F)$ есть столбец j-й матрицы производных A, поэтому можно записать:

$$G = \frac{1}{2} \left\{ \frac{\partial \varphi}{\partial x_j} \right\} = A^T F; \quad j = \overline{1,n} \tag{1.19}$$

Выражение для градиента G_L линейной модели φ_L, получим, заменив F в предыдущей формуле на линейную модель (1.16)

$$G_L = A^T (F_0 + A\Delta X) = A^T F_0 + A^T A \Delta X = N + M\Delta X = G_0 + M\Delta X \tag{1.20}$$

где N и M – ранее введенные матрицы, а $G_0 = N$ – градиент в начальной точке. Выражение (1.20) может быть также получено непосредственным дифференцированием выражения (1.17) для оценочной функции линейной модели.

Заметим, что в точке $X_{\min}$ минимума оценочной функции градиент G с необходимостью равен нулю.

$$G(X_{\min}) = 0 \tag{1.21}$$

1.5. ПОСТРОЕНИЕ ТРАЕКТОРИИ ДВИЖЕНИЯ.

Выбор той или иной траектории движения является основной чертой, определяющей эффективность оптимизации, отличающей различные методы и обычно определяющей название метода. Построить траекторию – значит указать конкретный вид зависимости вектора $\Delta X(s)$ точки на этой

траектории от скалярного параметра s, называемого длиной шага. Вектор $\Delta X(s)$ есть вектор отклонения точки на траектории X от начальной точки X_0 данного шага оптимизации. Предполагается, что $\Delta X(0)=0$, т.е. длина шага, равная нулю, соответствует начальной точке X_0; при увеличении s мы удаляемся по траектории от начальной точки. Рабочая ветвь траектории соответствует положительным значениям s.

Рассмотрим траектории для нескольких употребительных методов. В **градиентном методе** (сокращенно обозначим его ГМ), называемом также методом быстрейшего спуска [3, 6, 8], траектория движения представляет собой прямую линию, противоположную по направлению градиенту в начальной точке (рис. 1.3),

$$\Delta X(\mathrm{s}) = -\,\mathrm{s}G_0 = -\,\mathrm{s}N = -\,\mathrm{s}A^T F_0 \qquad (1.22)$$

Найдем значение длины шага s, соответствующее оптимальной точке на траектории в рамках линейной модели, т.е. той точке, в которой оценочная функция φ_L линейной модели достигает минимума. Для этого подставим уравнение (1.22) траектории в выражение (1.17) для оценочной функции линейной модели

$$\varphi_L(s) = \varphi_0 - 2sN^T N + s^2 N^T MN \qquad (1.23)$$

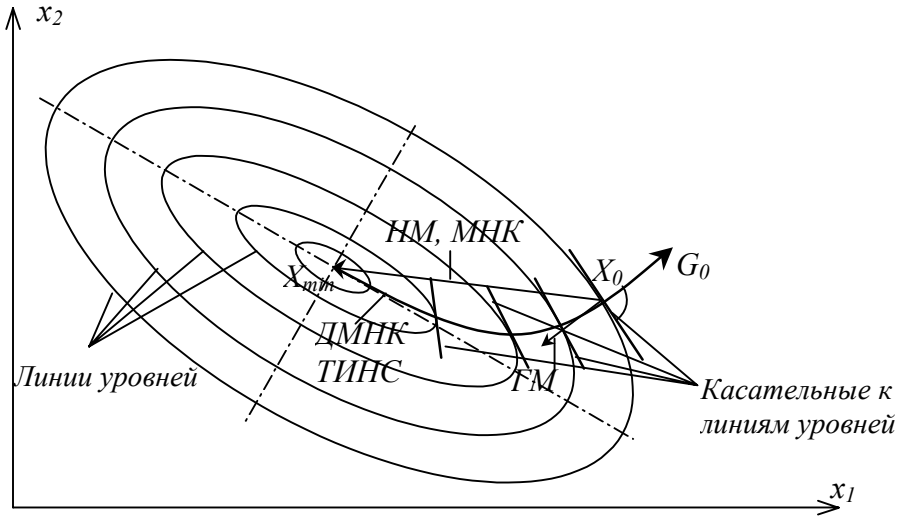

Рис.1.3.

Дифференцируя это выражение по s и приравнивая производную нулю, получаем искомый минимум $s_{\min}$ на траектории

$$s_{\min} = \frac{N^T N}{N^T MN} \qquad (1.24)$$

В том, что полученное значение является минимумом, легко убедиться найдя вторую производную $\varphi_L(s)$ по s

$$\frac{\partial^2 \varphi_L(s)}{\partial s^2} = N^T M N$$

Так как M – неотрицательно определенная матрица, то предыдущее выражение есть неотрицательно определенная квадратичная форма N, т.е. при любых N это выражение ≥ 0, следовательно, функция $\varphi_L(s)$ имеет один экстремум, являющийся минимумом. (Выражение (1.23) показывает, что график этой функции – парабола, вершина которой является минимумом.) Полезно заметить, что в точке $s_{\min}$ на траектории, т.е. в точке $\Delta X(s_{\min})$, градиент ортогонален траектории, или, что то же самое, ортогонален градиенту G_0 в начальной точке. Действительно, ортогональность означает, что скалярное произведение соответствующих векторов равно нулю. В матричной форме скалярное произведение записывается как произведение транспонированной матрицы-столбца одного вектора на матрицу-столбец другого. Обозначив скалярное произведение векторов G_0 и G через $(G_0 \cdot G)$, запишем

$$(G_0 \cdot G) = G_0^T (N + M\Delta X) = N^T (N - s_{\min} M G_0) =$$

$$= N^T \left(N - \frac{N^T N}{N^T M N} \right) = N^T N - \frac{N^T N}{N^T M N} N^T M N = 0$$

Представим теперь длину шага s в виде произведения $s = \rho s_{\min}$, где $s_{\min}$ длина шага, соответствующая минимуму линейной модели на траектории, определяемая выражением (1.24), ρ – безразмерная относительная длина шага. Тогда траекторию в градиентном методе можно представить в виде

$$\Delta X(\rho) = -\rho s_{\min} N, \text{ где } s_{\min} = \frac{N^T N}{N^T M N} \qquad (1.25)$$

Смысл описанного выбора траектории очевиден. Если градиент $G_0 = N$ указывает направление быстрейшего возрастания оценочной функции φ, а целью оптимизации является насколько возможно уменьшение последней, то движение от исходной точки в направлении, противоположном градиенту, дает наибыстрейшее из всех других возможных направлений убывание, что и требуется. Благодаря такому очевидному соответствию задаче оптимизации, а также крайней простоте, указанный метод приобрел большую популярность. Легко, однако, убедиться, что траектория (1.25) в общем случае не проходит через минимум линейной модели, поэтому для оптимизации систем, даже близких к линейным, метод часто требует громадного количества шагов и, следовательно, такого же количества проб и проб производных, поэтому при сколько-нибудь значительной трудоемкости последних он может быть совершенно непригоден. На эффективность метода в сильнейшей степени влияет выбор метрики M_x пространства параметров. Более подробно эти вопросы будут рассмотрены в следующем параграфе.

Модифицированный вариант градиентного метода носит название **метода сопряженных градиентов** (СГМ) или параллельных касательных [6, 9].

Первый шаг процесса в этом методе производится точно так же, как и в ГМ, а направление траектории второго и последующих шагов корректируется с учетом предыдущего шага

$$\Delta X(\rho) = \rho s_{\min} D^{(k)}, \tag{1.26}$$

где

$$D^{(k)} = -\left\{ N^{(k)} - \frac{\left[N^{(k)}\right]^T \left[N^{(k)}\right]}{\left[N^{(k-1)}\right]^T \left[N^{K-1}\right]} D^{(k-1)} \right\};$$

причем $D^{(0)} = -G_0^{(0)} = -N^{(0)}$, а $s_{\min}$ выбирается из тех же соображений, что и в ГМ. Проделав аналогичные рассуждения, получим

$$s_{\min} = \frac{N^T N}{D^T M D} \tag{1.27}$$

Можно убедиться в том, что не более чем на n шаге траектория СГМ пройдет через минимум линейной модели.

Таким образом, метод сопряженных градиентов более эффективен для систем, близких к линейным, чем градиентный метод, и сохраняет тем не менее простоту последнего.

Стремление как можно более полно использовать информацию, заключенную в матрице производных A с тем, чтобы выбранная траектория приводила к минимуму линейной модели за один шаг, вызывает необходимость решения тех или иных систем линейных уравнений, полученных из матрицы A.

Метод Ньютона (НМ), называемый также методом касательных или методом линеаризованных итераций, может применяться только в тех случаях, когда размерности пространств параметров и функций совпадают, т.е. когда $m = n$. Этот метод основан на решении линейной системы, полученной из приравнивания линейной модели (1.16) нулю

$$F_L = F_0 + A\Delta X = 0 \text{ или } A\Delta X = -F_0 \tag{1.28}$$

При $m = n$ матрица A предыдущего уравнения квадратная и в принципе это уравнение может быть решено относительно ΔX. Решение этого уравнения (записанное символически в виде $\Delta X = -A^{-1}F_0$, где A^{-1} матрица, обратная A), умноженное на безразмерную длину шага ρ, и дает траекторию движения в этом методе

$$\Delta X(\rho) = -\rho A^{-1} F_0 \tag{1.29}$$

Видно, что траектория представляет собой прямую линию, проходящую через минимум (обязательно нулевой) линейной модели (рис. 1.3). Очевидно, что для систем, близких к линейным, минимум по методу Ньютона может быть найден за небольшое количество шагов, в этом смысле НМ является

более эффективным, чем предыдущие, но одновременно он и более трудоемок, так как для определения траектории требует решения линейной системы, связанного с выполнением порядка n^3 действий, Кроме того, этот метод может расходиться, т.е. решение (1.29) может не существовать, или оно становится недопустимо большим, если среди строк или столбцов матрицы имеются линейно зависимые, что для больших систем является нередким.

Развитием метода Ньютона на случай $m > n$ можно считать известный **метод наименьших квадратов** (МНК.) [9]. Траектория движения в этом методе определяется из решения так называемой нормальной системы, полученной приравниванием нулю градиента G_i линейной модели.

$$G_L = N + M\Delta X = 0 \text{ или } M\Delta X = -N \qquad (1.30)$$

Решение этой системы $\Delta X = -M^{-1}N$ определяет точку, в которой градиент G_L равен нулю, т.е. точку минимума для линейной модели. Умножив это решение на безразмерную длину шага ρ получим траекторию движения $\Delta X(\rho)$ в виде прямой линии, проходящей при $\rho = 1$ через минимум линейной модели,

$$\Delta X(\rho) = -\rho M^{-1}N = -\rho(A^T A)^{-1}A^T F_0 \qquad (1.31)$$

Полезно рассмотреть некоторые свойства этого широко распространенного метода. Покажем, что при $m = n$ метод идентичен (до погрешностей вычислений) методу Ньютона. Действительно, подставляя в (1.31) выражения для M и N и учитывая, что при $m = n$ матрицы A и A^T квадратные, следовательно, для них могут существовать обратные матрицы, получим

$$\Delta X(\rho) = -\rho(A^T A)^{-1}A^T F_0 = -\rho A^{-1}(A^T)^{-1}A^T F_0 = -\rho A^{-1}F_0$$

что тождественно (1.29). При выводе предыдущего выражения мы воспользовались следующими правилами матричной алгебры: $(AB)^{-1} = B^{-1}A^{-1}$; $A^{-1}A = AA^{-1} = I$ и $B \cdot I = B$; I – единичная матрица; A и B – произвольны квадратные неособенные матрицы.

Исследуем теперь существование решения (1.31) в общем случае. Известно, что если матрица M – особенная (вырожденная), то обратная матрица M^{-1} не существует и решение (1.31) отсутствует (становится бесконечно большим). Для того чтобы проверить, не является ли матрица M вырожденной, необходимо знать ее ранг r_M т. е. максимальный порядок определителя, составленного из строк и столбцов этой матрицы, не равного нулю. Максимально возможный ранг квадратной матрицы, очевидно, равен ее порядку, в этом случае матрица не вырождена. Если ранг матрицы меньше ее порядка, матрица вырождена. Найдем ранг матрицы M. Как можно показать, ее ранг равен рангу матрицы A [7]:

$$r_M = r_A$$

Пусть в матрице производных A имеется $t+1$ линейно зависимых строк и $q+1$ линейно зависимых столбцов. Это означает, что из m функций t представляют собой линейную комбинацию остальных и, аналогично, из n параметров q есть линейная комбинация остальных. Это эквивалентно тому, что фактически в системе вместо m функций и n параметров имеется только $m-t$ независимых функций и $n-q$ независимых действенных параметров. Действительно, эквивалентными преобразованиями, не нарушающими ранг матрицы (т.е. перестановкой двух строк или столбцов, умножением какой-либо строки или столбца на любое число, отличное от нуля, и прибавлением к какой-либо строке или столбцу другой строки или столбца, умноженных на отличное от нуля число), можно привести матрицу A к матрице, имеющей t нулевых строк и q нулевых столбцов. Тогда легко убедиться в том, что ранг матрицы A и, следовательно, ранг матрицы M не превосходит меньшего из чисел $m-t$ и $n-q$, которые являются числами ненулевых строк и столбцов в преобразованной матрице A

$$r_M = r_A \le \min(m-t, n-q) \tag{1.32}$$

Матрица M вырождена, если ее ранг меньше порядка, т.е.

$$\min(m-t, n-q) < n \tag{1.33}$$

Пусть $m < n$, тогда $\min(m-t, n-q) < n$ при любом t, следовательно в этом случае МНК не работает.

Пусть теперь $m > n$. Условие вырожденности (1.33), т.е. отсутствия решения по МНК, выполняется, если

$$q > 0 \tag{1.34}$$

или

$$t > m-n\,. \tag{1.35}$$

Таким образом, МНК не имеет решения (расходится), если среди параметров имеется хотя бы один недействующий ($q>0$) (или два линейно связанных) или среди функций имеется хотя бы $m-n+1$ не зависящих ни от одного параметра ($m-n+2$ линейнозависимых между собой). Очевидно, что вероятность выполнения одного из условий (1.34) и (1.35) тем меньше, чем больше в системе функций и чем меньше параметров, т.е. чем больше отношение m/n при прочих равных условиях. Практика показывает, что МНК более или менее устойчиво работает при $m/n > 10$ [10].

С целью обеспечения устойчивости МНК во всех случаях, последний был усовершенствован, получив название **демпфированного метода наименьших квадратов** (ДМНК) [12]. Сущность этого радикального усовершенствования заключается в так называемом демпфировании, т.е. добавлении в оценочную функцию φ взвешенной суммы квадратов отклонений параметров от их значений в начальной точке, т.е. квадрата длины вектора

$$\varphi_p = \varphi + p^2(\Delta X)^2 = \varphi + p^2 \Delta X^T \Delta X = F^T F + p^2 \Delta X^T \Delta X \qquad (1.36)$$

где φ_p – новая оценочная функция; p^2 – весовой коэффициент, демпфер. В процессе оптимизации ищется уже не минимум φ, а минимум φ_p, т.е. к условию минимума квадрата длины вектора функций присоединяется требование минимизации квадрата длины отклонения ΔX от начальной точки в пространстве параметров. Определим точку, в которой демпфированная оценочная функция φ_{pL} линейной модели имеет минимум

$$\begin{aligned} \varphi_{pL} &= \varphi_L + p^2 \Delta X^T \Delta X = \varphi_0 + \Delta X^T N + \Delta X^T M \Delta X + \\ &+ p^2 \Delta X^T \Delta X = \varphi_0 + \Delta X^T N + \Delta X^T \left(M + p^2 I\right) \Delta X \end{aligned} \qquad (1.37)$$

Для этого найдем градиент G_{pL} этой оценочной функции и приравняем его нулю

$$G_{pL} = N + (M + p^2 I)\Delta X = 0 \qquad (1.38)$$

или

$$M_p \Delta X = -N, \text{ где } M_p = M + p^2 I \qquad (1.39)$$

В результате получим так называемую демпфированную нормальную систему уравнений относительно ΔX, матрица M_p которой отличается от матрицы M нормальной системы в МНК тем, что к ней добавлена умноженная на демпфер p^2 единичная матрица или, другими словами, тем, что диагональные элементы матрицы M увеличены на p^2.

Как и матрица M, матрица M_p симметрическая. Решение системы (1.39) может быть записано в виде

$$\Delta X = -M_p^{-1} N = -(M + p^2 I)^{-1} N \qquad (1.40)$$

Покажем, что при любом отличном от нуля демпфере матрица M_p невырождена, т. е. решение (1.40) всегда существует и конечно.

Для этого рассмотрим спектр так называемых собственных значений матриц M и M_p [2, 7]. Так как матрица M – неотрицательно определенная, то ее собственные значения действительны и неотрицательны. Можно показать (см. §1-9), что собственные значения матрицы M_p равны собственным значениям матрицы M, увеличенным на демпфер p^2, т. е. при любом $p^2 > 0$ всегда положительны и тем больше, чем больше p. При возрастании p все собственные значения матрицы M_p приближаются к p^2.

С другой стороны, известно, что матрица вырождена, если среди ее собственных значений имеются нулевые. Следовательно, матрица M_p при любом ненулевом p невырождена. Более того, можно рассматривать так

называемую степень обусловленности матрицы M_p, которая показывает устойчивость решения системы, а именно, чем больше степень обусловленности M_p, тем менее устойчиво решение. Степень обусловленности равна отношению максимального и минимального собственных значений матрицы, отсюда видно, что с увеличением p^2 степень обусловленности матрицы M_p уменьшается, стремясь в пределе к единице, т.е. матрица становится все лучше и лучше обусловленной.

Рассмотрим поведение решения (1.40) при различных значениях демпфера p. Очевидно, что при малых p, в пределе при $p=0$, решение (1.40) переходит в решение (1.31) по-обычному МНК. При возрастании p^2, в пределе при $p^2 \to \infty$, в матрице M_p член $p^2 I$ преобладает над M, которым поэтому можно пренебречь. Таким образом, при больших p^2 имеем

$$\Delta X = -(p^2 I)^{-1} N = -\frac{1}{p^2} N \qquad (1.41)$$

Видно, что в этом случае ДМНК дает решение, пропорциональное градиенту. Таким образом, при малых p^2 решение направлено на минимум линейной модели, а при больших p^2 – в сторону, противоположную градиенту.

Рассмотрим теперь уравнение траектории движения в ДМНК. Это уравнение может определяться двояким образом, В первом случае траектория движения представляет собой прямую линию

$$\Delta X(\rho) = -\rho(M + p^2 I)^{-1} N \qquad (1.42)$$

где демпферу p^2 присваивается какое-либо фиксированное значение.

Во втором случае траектория есть кривая линия

$$\Delta X(\rho) = -[M + p^2(\rho) I]^{-1} N \qquad (1.43)$$

где $p^2(\rho)$ – некоторая функция, показывающая зависимость демпфера от безразмерной длины шага ρ. Из условия удаления по траектории с увеличением ρ следует, что $p^2(\rho)$ должна быть монотонно убывающей функцией, равной бесконечности при $\rho = 0$. Выбор этой функции, от которой существенно зависит эффективность оптимизации, требует дополнительного исследования, в качестве первого приближения можно принять функцию вида: $p^2(\rho) = \frac{p_0^2}{\rho^2}$, где p_0 – некоторое начальное значение демпфера.

Выбор траектории в качестве прямой линии несомненно проще, так как здесь требуется только один раз на данном шаге решить линейную систему

(1.39), а движение вдоль траектории получается простым умножением вектора ΔX решения системы (1.39) на длину шага ρ.

Во втором случае при движении вдоль траектории для каждого нового значения длины шага ρ необходимо решать линейную систему (1.39). Поскольку решение линейной системы порядка n требует около n^3 действий, движение по криволинейной траектории значительно более трудоемко, особенно при больших n. Следует ожидать, однако, что эта трудоемкость оправдывается большей эффективностью оптимизации. Действительно, как показали исследования [11], движение вдоль кривой линии с изменением демпфера в соответствии с (1.43) гораздо эффективнее, чем движение по прямой линии (1.42) с фиксированным значением демпфера, кривая линия (1.43) почти всегда проходит вблизи минимума системы, в то время как прямая приближается к минимуму только при определенном значении демпфера. Следовательно, при движении по кривой достаточно одного параметра ρ (выбор которого будет рассмотрен в следующем параграфе), чтобы подойти близко к минимуму, при движении по прямой приходится выбирать два параметра ρ и p, что требует более сложного алгоритма и в конце концов приводит к большой затрате проб. Вследствие указанных причин выбор траектории в виде кривой линии является основной рабочей модификацией ДМНК.

1.6. ВЫБОР ДЛИНЫ ШАГА ВДОЛЬ ТРАЕКТОРИИ

Выбор длины шага заключается в определении конкретного значения ρ, которое, будучи подставлено в уравнение траектории, дает вектор $X_l = X_0 + \Delta X(\rho)$ – конечной точки данного шага и начальной точки следующего шага. Этот выбор должен обеспечивать сходимость метода, т.е. выполнение условия $\varphi(X_l) < \varphi(X_0)$ или, учитывая, что X есть функция от ρ, описываемая уравнением траектории, условие сходимости может быть записано в виде

$$\varphi(\rho) < \varphi(0) = \varphi_0 \qquad (1.44)$$

Кроме того, желательно, чтобы выбор ρ обеспечивал наилучшую сходимость, позволяя продвинуться к искомому минимуму как можно ближе, насколько позволяет данная траектория.

Для систем, близких к линейным, наилучшее значение ρ близко к единице, как это следует из предыдущего параграфа. Однако для нелинейных систем такой выбор может привести к расходимости метода, т.е. невыполнению условия (1.44).

Рассмотрим некоторые употребительные методы выбора ρ.

Метод релаксации [10] предполагает выбор постоянной для всех шагов и малой длины шага $\rho << 1$, с тем чтобы для любых оптимизируемых систем в пределах выбранной малой ρ была заранее гарантирована справедливость линейной модели. Легко показать, что для всех рассмотренных методов в

области линейной модели выбор $\rho < 1$ удовлетворяет условию сходимости. Этот выбор, однако, не является удачным, поскольку необходимость выбора очень малых значений ρ при оптимизации нелинейных систем приводит к очень медленной сходимости и, следовательно, большой цене поиска.

Метод слежения за областью линейности [11] предусматривает корректировку длины ρ шага на каждой итерации, с тем чтобы продвинуться вдоль траектории как можно дальше в пределах области линейности. Показателем отступления реальной системы от линейной может служить величина

$$\theta = \frac{\varphi - \varphi_0}{\varphi_L - \varphi_0}$$

где φ_0 − значение оценочной функции в начальной точке данного шага; φ_L − значение оценочной функции линейной модели для данной ρ, определенное по формуле (1-17); φ — значение оценочной функции реальной системы для данной ρ, определенное из пробы в точке $X_0 + \Delta X(\rho)$.

Если $0.9 > \theta > 0.5$, то система может считаться близкой к линейной и значение ρ, полученное на данном k-м шаге, сохраняется и для $(k+1)$-го шага: $\rho_{k+1} = \rho_k$.

Если $\theta > 0.9$, то система остается линейной и при большой ρ, поэтому есть смысл на следующем шаге увеличить ρ: $\rho_{k+1} = \beta\rho_K$ (β − некоторый положительный коэффициент, обычно $\beta = 2 \div 5$). Если $0 > \theta > 0.5$, то для выбранного ρ система выходит за область линейности, поэтому на следующем шаге рационально уменьшить ρ: $\rho_{k+1} = \frac{1}{\beta}\rho_k$.

Если $\theta < 0$, то система настолько нелинейна, что при выбранной ρ метод расходится уже на данном шаге, т.е, не выполняется условие сходимости (1.44). В этом случае ρ перевычисляется уже для данного шага $\rho_k = \frac{1}{\beta}\rho_k$ и в новой точке повторяется проба.

Рассмотренный метод выбора ρ практически не требует дополнительных затрат проб и в то же время весьма эффективен,

Наиболее эффективной, однако, оказывается **оптимизация оценочной функции** по длине шага, т.е. выбор такого $\rho = \rho_{\min}$, для которого $\varphi(\rho)$ имеет минимум. Так как оптимизация здесь производится по одному параметру, то наиболее эффективны глобальные однопараметрические методы оптимизации. Рассмотрим, например, **метод параболической аппроксимации** [5]. Сначала при помощи нескольких проб на траектории находим три такие точки, для которых выполняется условие

$$\varphi_1 = \varphi(\rho_1) > \varphi_2 = \varphi(\rho_2) < \varphi_3 = \varphi(\rho_3)$$

Затем через эти три точки проводится парабола, и ее вершина принимается за искомый минимум:

$$\rho_{\min} = \frac{1}{2}\frac{(\varphi_1 - \varphi_2)\rho_3^2 + (\varphi_2 - \varphi_3)\rho_1^2 + (\varphi_3 - \varphi_1)\rho_2^2}{(\varphi_1 - \varphi_2)\rho_3 + (\varphi_2 - \varphi_3)\rho_1 + (\varphi_3 - \varphi_1)\rho_2} \qquad (1.46)$$

Методы, использующие такой выбор длины шага, называются оптимальными. Хотя они несколько более трудоемки, поскольку для нахождения $\rho_{\min}$ требуется затратить несколько проб, (обычно не более 4–5), эти затраты в итоге компенсируются большей эффективностью движения на каждом шаге и, как следствие, уменьшением потребного количества шагов.

1.7. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПОСТРОЕНИЯ ТРАЕКТОРИИ

Выбор траектории движения, рассмотренный в п.1.5, определяющий название метода оптимизации, определяет и успех оптимизации, ее скорость и потраченное время. Поэтому рассматриваемый в настоящем параграфе сравнительный анализ различных методов (при неизбежной его краткости) может быть весьма полезен в практической работе по оптимизации.

Мы будем исходить при оценке методов из двух критериев: универсальности метода и его эффективности при оптимизации.

Вопрос с первым критерием решается весьма просто. В п.1.5 при рассмотрении различных методов мы пришли к выводу, что ГМ, СГМ и ДМНК дают решение для любых систем, т.е. универсальны, в то время как НМ и МНК не обладают таким свойством. НМ может применяться только для систем, у которых количество параметров равно количеству функций, МНК – для систем, у которых количество функций превышает количество параметров, Кроме того, эти методы не работают, т.е. не дают конечного решения, если среди параметров имеется хотя бы один невлияющий или два линейно связанных, или если среди функций имеются хотя бы $m-n+1$ не зависящих ни от одного из параметров ($m-n+2$ линейно зависимых).

Решение вопроса об эффективности того или иного метода более затруднительно. В п.1.1 мы договорились определять эффективность оптимизации так называемой ценой поиска, т.е. количеством проб, затраченных в процессе поиска. Если учесть в цене поиска еще и трудоемкость самого метода, т.е. необходимое для решения количество вычислений, то, исходя из материала предыдущих параграфов, можно написать следующее выражение для цены поиска:

$$C = k[c_1 + c_2 + l(1 + c_3)]$$

где C – цена поиска; k – количество шагов поиска, необходимое для достижения минимума с заданной точностью; l – количество проб, затраченных на однопараметрическую оптимизацию, т.е. на выбор длины шага $\rho_{\min}$; c_2 – сравнительная трудоемкость построения траектории, выраженная в пробах; c_3 – сравнительная трудоемкость движения по

траектории, т.е. определения вектора точки на траектории по заданной ρ, выраженная в пробах; c_1 – сравнительная трудоемкость пробы производных, выраженная в пробах. (В большинстве случаев вычисление производных осуществляется методом конечных разностей, поэтому можно принять $c_1 = n$).

Все величины, входящие в предыдущее выражение, могут быть легко определены для любого метода. Так, в зависимости от способа выбора длины шага ρ количество пробных точек на траектории может колебаться от нуля при методе релаксации и 0.5 (в среднем) для метода слежения за областью линейности до 4–5 для оптимального метода. Сравнительная трудоемкость построения траектории и движения по траектории c_2 и c_3 у ГМ и СГМ практически может быть принята равной нулю, у НМ и МНК c_2 уже сравнима с единицей, так как для построения траектории здесь требуется решение линейной системы порядка n, т.е. выполнение $\sim n^3$ действий. Такое количество действий, как показывает статистика использования оптимизаторов, характерно для проб средней трудоемкости. Трудоемкость движения по траектории у этих методов может быть принята равной нулю (умножение на ρ). ДМНК — самый трудоемкий метод, у него c_2 и c_3 сравнимы с единицей. Таким образом, в пределах одного шага нетрудно оценить относительную трудоемкость различных методов, и, как следует из изложенного, для средних проб она будет колебаться от $n+5$ для ГМ до $n+11$ для ДМНК, т.е. при больших $n >> 1$ трудоемкость различных методов практически одинакова в пределах шага.

Следовательно, единственным параметром в формуле (1.47), определяющим различную эффективность различных методов, может быть *k* – количество шагов оптимизации,

К сожалению, сколько-нибудь строгий анализ потребного количества шагов, не зависящий от априорных сведений об оптимизируемой системе, невозможен. При этом естественно, что *k* в сильнейшей степени зависит от характера нелинейности оптимизируемой системы и от выбора начальной точки, более того, можно показать, что всегда возможно сконструировать такую систему и взять такую начальную точку, для которых любой из произвольно выбранных методов будет самым эффективным. Поэтому дальнейшие рассуждения определяют только вероятность эффективности того или иного метода, в конкретных случаях поведение методов может, вообще говоря, не соответствовать полученным оценкам.

Все рассмотренные методы основаны на линейной модели оптимизируемой системы в окрестности начальной точки каждого шага, т.е. используют для построения траектории информацию только линейной модели. Исходя из этого, можно оценивать методы по эффективности использования этой информации. Введем понятие об области линейности системы на каждом шаге. Это область вокруг начальной точки каждого шага, в которой поведение оценочной функции реальной системы незначительно

отличается от поведения оценочной функции линейной модели, описываемого выражением (1.17). Другими словами, нелинейные добавки к оценочной функции линейной модели в области линейности должны быть малы по сравнению с членами $\Delta X^T N$ и $\Delta X^T M \Delta X$ выражения (1.17). Пусть φ – оценочная функция реальной системы; φ_L – оценочная функция линейной модели; $\Delta\varphi$ – нелинейные добавки в оценочную функцию

$$\Delta\varphi = \varphi - \varphi_L$$

Область линейности L есть область, удовлетворяющая условию

$$\left|\frac{\Delta\varphi}{2\Delta X^T N + \Delta X^T M \Delta X}\right| << 1 \text{ для всех } \Delta X \in L .$$

Наиболее эффективный метод, основанный на линейной модели, должен за каждый шаг обеспечивать продвижение к минимуму на расстояние, примерно равное радиусу r_L области линейности, пока искомый минимум не попадет в очередную область, после чего на этом шаге он должен быть найден с требуемой точностью (может быть, потребуется еще один шаг для уточнения, если допускаемая погрешность отыскания минимума меньше погрешности в определении области линейности). Для такого метода количество шагов k определяется отношением расстояния от исходной точки до минимума к радиусу области линейности (среднему).

Легко убедиться, что из рассмотренных методов ГМ и СГМ не являются эффективными в этом смысле. Действительно, даже для строго линейных систем, у которых областью линейности является все пространство параметров, траектория в этих методах в общем случае не проходит через минимум линейной модели.

Из рис.1.3 видно, что только в тех частных случаях, когда начальная точка данного шага близка к одной из четырех вершин эллипсоида поверхности уровня оценочной функции линейной модели, траектория ГМ проходит вблизи минимума. Так как площадь этих благоприятных малых областей в окрестности вершин пренебрежимо мала по сравнению со всей площадью гиперэллипсоида, то вероятность удачного выбора начальной точки практически равна нулю для отличающегося от сферы эллипсоида. Количество шагов, за которое траектория ГМ приближается к минимуму линейной модели с заданной точностью, зависит от формы эллипсоида и, как показывает анализ, в практических случаях может достигать величин $10^2 - 10^3$ и более. Траектория СГМ точно проходит через минимум линейной модели не более чем на n-м шаге, где n – количество параметров, что при больших n $(n > 10)$ не намного лучше ГМ. Таким образом, в подавляющем большинстве случаев сходимость методов ГМ и СГМ крайне низкая и не увеличивается при приближении к минимуму, когда последний попадает в область линейности и линейная модель содержит достаточно информации о его положении.

Можно поэтому с уверенностью утверждать, что, несмотря на их простоту, эти методы не могут быть положены в основу оптимизирующей программы.

Рассмотрим теперь метод наименьших квадратов (МНК). Ранее было показано, что в случае, если этот метод имеет решение, то его траектория проходит через минимум линейной модели, следовательно, можно ожидать, что этот метод будет близок к эффективному.

Рассмотренные выше условия отсутствия решения по этому методу на практике почти никогда строго не выполняются, поскольку наличие в системе строго не влияющих или линейно зависимых параметров мало вероятно. В реальных оптимизируемых нелинейных системах более вероятно появление в окрестности той или иной начальной точки слабо влияющих параметров или параметров, близких к линейно связанным, причем эти свойства параметров сохраняются в небольшой области, далее эти параметры становятся влияющими, а другие – слабо влияющими и т. д.

Рассмотрим изменение траектории метода при уменьшении влияния какого-либо параметра до нуля, т.е. при приближении матрицы M к вырожденной.

Изменение влияния параметров эквивалентно умножению столбцов матрицы производных A, соответствующих различным параметрам, на относительные коэффициенты влияния γ_j, т.е. может быть описано как умножение матрицы производных A справа на диагональную матрицу влияния Γ, диагональные элементы которой равны коэффициентам относительного влияния различных параметров

$$A_\gamma = A\Gamma \tag{1.47}$$

где A – матрица производных по старым параметрам; A_γ – матрица производных по параметрам с измененным влиянием; Γ – диагональная матрица влияния;

$$\Gamma = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix}.$$

Найдем новое решение системы (1.30) с учетом влияния Γ. Из (1.30) и (1.47) получим

$$\begin{aligned} \Delta X_\gamma &= (A_\gamma^T A_\gamma)^{-1} A_\gamma^T F_0 = [(A\Gamma)^T A\Gamma]^{-1} (A\Gamma)^T F_0 = \\ &= \left(\Gamma^T A^T A\Gamma\right)^{-1} \Gamma^T A^T F_0 = \Gamma^{-1} M^{-1} \Gamma^{-1} \Gamma A^T F_0 = \\ &= \Gamma^{-1} M^{-1} N = \Gamma^{-1} \Delta X \end{aligned} \tag{1.48}$$

где ΔX_γ – вектор решения системы с измененным влиянием параметров; ΔX – прежний вектор решения системы. Из предыдущего

выражения видно, что в МНК при изменении влияния параметров соответствующие проекции вектора решения делятся на коэффициенты влияния: $\Delta X_{\gamma j} = \dfrac{\Delta X_j}{\gamma_j}$, т. е. если влияние какого-либо параметра стремится к нулю, соответствующая ему проекция вектора стремится к бесконечности. При подстановке решения (1.48) в выражение для траектории (1.31), мы приходим к необходимости выбора в процессе движения по траектории крайне малых величин длины шага ρ, чтобы изменение слабо влияющего параметра не вышло из области его линейности. В результате изменение остальных, сильно влияющих параметров, получается ничтожно малым и сходимость метода сильно тормозится.

Таким образом, в случаях большой разницы во влиянии различных параметров на оптимизируемую систему, при наличии слабо влияющих хотя бы на данном шаге параметров метод наименьших квадратов теряет свою эффективность.

Наиболее рационально построенный метод должен, наоборот, обеспечивать стремление приращения слабо влияющего параметра к нулю, тем самым автоматически исключая его из оптимизации и нейтрализуя его тормозящее влияние на сходимость. Заметим, что градиентный метод удовлетворяет этому условию, что следует из (1.25) и (1.47). Подставляя в выражение (1.22) для траектории ГМ измененную матрицу производных $A_\gamma = A\Gamma$, получим

$$\Delta X_\gamma = -s(A\Gamma)^T F_0 = -s\Gamma A^T F_0 = \Gamma \Delta X, \qquad (1.49)$$

т. е. $\Delta x_{\gamma j} = \gamma_j \Delta x_j$ при $\gamma_j \to 0$ и $\Delta X_j \to 0$.

Следовательно, эффективный метод построения траектории должен объединять в себе положительные свойства ГМ и МНК. Именно такими свойствами обладает демпфированный метод наименьших квадратов с движением по кривой.

Действительно, как было показано ранее, этот метод универсален, он имеет решение при любом соотношении количества параметров и функций, независимо от того, имеются среди них линейно связанные или не влияющие параметры или нет.

При малых длинах шага ρ направление траектории этого метода приближается к траектории ГМ, а при больших ρ траектория ДМНК приходит к минимуму линейной модели.

Следует сказать, что существует распространенное мнение о высокой эффективности градиентного метода (метода быстрейшего спуска). Исследованию этого метода посвящена многочисленная литература. Указанное заблуждение основано на неправильном отождествлении траектории ГМ с криволинейной траекторией истинного наибыстрейшего спуска (ТИНС), которая представляет собой линию, ортогональную всем поверхностям уровня. В действительности ГМ есть лишь касательная к

ТИНС в начальной точке и очень быстро отклоняется от нее по мере продвижения по траектории. Покажем, что траектория ТИНС совпадает с траекторией ДМНК. Для этого необходимо рассмотреть сферу радиуса R с центром в начальной точке X_0 и определить точку на указанной сфере, в которой оценочная функция φ была минимальна.

Таким образом, задача сводится к определению минимума φ на сфере радиуса R

$$\varphi = F^T F - \min;$$

$$\Delta X^T \Delta X = R^2.$$

Указанная задача определения условного минимума может быть сведена с помощью множителей Лагранжа к следующему:

$$\Phi = F^T F + \lambda\left(\Delta X^T \Delta X - R^2\right) - \min;$$

$$\Delta X^T \Delta X - R^2 = 0.$$

Последнее выражение дает решение в виде:

$$\Phi'_x = 0;$$

$$\Delta X^T \Delta X - R^2 = 0.$$

Продифференцировав Φ, получим:

$$2A^T F + 2\lambda \Delta X = 0;$$

$$\Delta X^T \Delta X = R^2.$$

Имея в виду, что $F_L = A\Delta X + F_0$, окончательно запишем:

$$\Delta X = -\left(A^T A + \lambda I\right)^{-1} A^T F_0;$$

$$\Delta X^T \Delta X = R^2.$$

Таким образом, определена точка $\Delta X = -\left(A^T A + \lambda I\right)^{-1} A^T F_0$, принадлежащая сфере радиуса R, в которой φ_L минимальна.

Легко увидеть, что через указанную точку сферы проходит траектория ДМНК при $p^2 = \lambda$.

Полученные результаты можно трактовать следующим образом.

При движении по траектории ДМНК каждому значению p^2 отвечает вектор ΔX из выражения (1.40). Этот вектор лежит на сфере радиуса $R = \Delta X^T \Delta X$, проведенной из точки X_0, причем значение оценочной функции φ_L, как было показано выше, в указанной точке $X_0 + \Delta X$ сферы наименьшее из всех точек выбранной сферы R. Другими словами ДМНК представляет собой траекторию самого эффективного движения в смысле уменьшения оценочной функции φ_L. Любая отличная от ДМНК траектория будет давать худшие результаты.

Поэтому ДМНК одинаково эффективен как в начале поиска, когда размеры области линейности малы по сравнению с расстоянием до искомого минимума, так и в конце, когда искомый минимум находится в области линейности.

Исследуем поведение метода при изменении влияния параметров, вводя матрицу влияния Γ. Подставив (1.47) в (1.48), получим

$$\Delta X_\gamma = -\left(\Gamma M \Gamma + p^2 I\right)^{-1} \Gamma N. \qquad (1.50)$$

В общем случае при изменении влияния параметров ΔX_γ меняется весьма сложным образом. Наиболее интересный случай, однако, соответствует стремлению влияния какого-либо j-го параметра к нулю, в этом случае матрица влияния Γ имеет вид

$$\Gamma = \begin{pmatrix} 1 & & & & & & & & 0 \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & \ddots & & & & & \\ & & & & \ddots & & & & \\ & & & & & \gamma_j & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & 1 \end{pmatrix}; \ \gamma_j \to 0. \qquad (1.51)$$

Тогда для приращения Δx_{γ_j}-го параметра можно получить приближенное выражение

$$\Delta x_{\gamma_j} \approx \Delta x_j \frac{\gamma_j}{\gamma_j^2 + p\left(1-\gamma_j^2\right)b_k}, \qquad (1.52)$$

где b_k – некоторый коэффициент, больший нуля. Отсюда видно, что при $\gamma_j \to 0$ и при любом p, отличном от нуля, приращение слабо влияющего параметра стремится к нулю, т.е. этот параметр исключается из оптимизации.

Имея в виду приведенные выше рассуждения, можно ожидать, что наибольшей эффективностью обладает ДМНК. Этот вывод подтверждается и практикой оптимизации. В следующем параграфе будет рассмотрен пример оптимизации несложной, но обладающей интересными свойствами системы, на котором ясно видны свойства каждого метода.

В заключение данного параграфа полезно заметить, что на эффективность различных методов оптимизации может существенно повлиять выбор метрики в пространстве параметров M_x.

В частности, в работах [8, 9] показано, что МНК и ДМНК могут быть сведены к градиентному методу изменением метрики M_x.

Пусть в пространстве параметров введена новая метрика таким образом, что

$$d_x^2 = X^T M_x X; \quad M_x = U\Lambda U^T.$$

Это означает переход от пространства X к пространству X_M, связанному с X линейным преобразованием

$$X_M = \Lambda^{\frac{1}{2}} U^T X \text{ или } X = U\Lambda^{-\frac{1}{2}} X_M. \tag{1.53}$$

Рассмотрим градиент G_M в пространстве с новой метрикой. Пусть известен градиент G в пространстве со старой метрикой. Так как проекции градиента G_M пропорциональны частным производным от φ по новым параметрам X_M, а градиента G – по старым параметрам X, легко получить соотношение между G_M и G

$$G_M = A_M^T F_0 = \left(AU\Lambda^{-\frac{1}{2}} \right)^T F_0 = \Lambda^{-\frac{1}{2}} U^T G = \Lambda^{-\frac{1}{2}} U^T N. \tag{1.54}$$

Таким образом, траектория ГМ в пространстве X_M будет иметь вид

$$\Delta X_M(s) = -sG_M = -s\Lambda^{-\frac{1}{2}} U^T N. \tag{1.55}$$

Перенесем эту траекторию обратно в пространство X, что соответствует умножению $\Delta X_M(s)$ на $U - \lambda^{1/2}$ в соответствии с (1.53)

$$\Delta X(s) = -sU\Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} U^T N = -sM_X N. \tag{1.56}$$

Таким образом, при изменении метрики пространства параметров траектория градиентного метода умножается слева на матрицу-метрику M_X. Сравнивая выражение (1.56) с выражениями (1.31) и (1.42) для траекторий МНК и ДМНК, мы видим, что выбором в пространстве параметров метрики в виде $M_X = (A^T A)^{-1}$ или $M_X = (A^T A + p^2 I)^{-1}$ ГМ можно перевести в МНК или ДМНК соответственно. Поэтому сравнительная эффективность методов зависит от выбора метрики пространства параметров.

Действительно, изменение метрики M_X, как было показано в п. 1.2, эквивалентно повороту системы координат и изменению масштаба по осям пространства параметров, поэтому соответствующим выбором метрики гиперэллипсоиды поверхностей уровней φ в пространстве параметров всегда можно деформировать в концентричные гиперсферы, для которых

градиент из любой точки направлен в центр, т. е. в минимум линейной модели.

Траектория МНК и ДМНК в меньшей степени зависит от изменения метрики пространства параметров, что является их преимуществом.

1.8. ПРИМЕР ОПТИМИЗАЦИИ

Для иллюстрации приведенных выше рассуждений рассмотрим пример оптимизации нелинейной системы из двух функций двух переменных:

$$\left.\begin{aligned} f_1 &= 10(x_1^2 - x_2); \\ f_2 &= 1 - x_1. \end{aligned}\right\} \qquad (1.57)$$

Примем, что пространства X и F уже нормированы. Выражение для оценочной функции этой системы имеет вид

$$\varphi = f_1^2 + f_2^2 = 100x_1^4 - x_1^2(200x_2 - 1) + 100x_2^2 - 2x_1 + 1. \qquad (1.58)$$

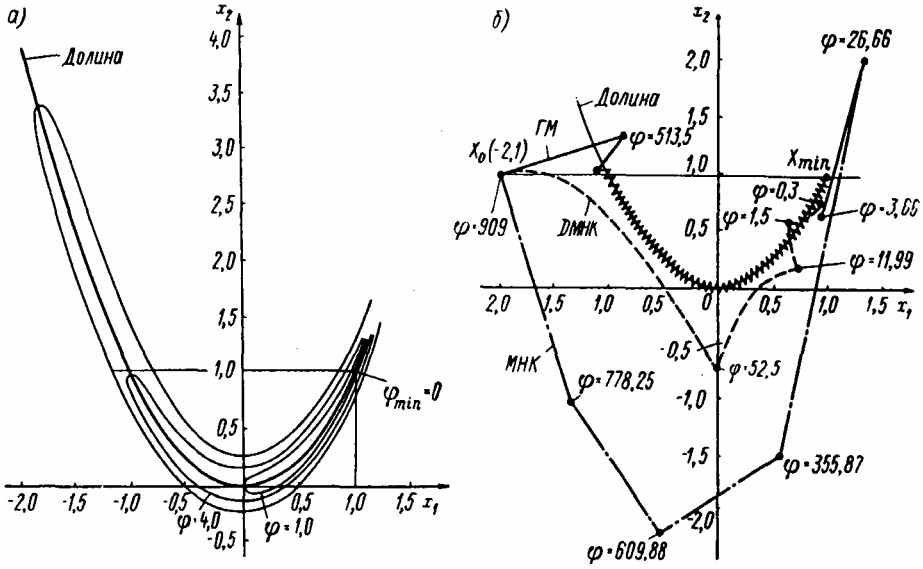

Рис. 1.4

На рис.1.4, *а* изображены линии уровня этой функции. Видно, что рельеф оценочной функции имеет ярко выраженную долину, центральная линия которой есть парабола, описываемая уравнением $x_2 = x_1^2$. Минимум оценочной функции (самая низкая точка долины) имеет координаты (1.1), которые для данного примера легко находятся аналитически из решения системы уравнений, полученной приравниванием (1.57) нулю.

В качестве исходной точки для всех методов была принята точка (-2;1). На рис.1.4, *б* изображены траектории движения к минимуму от исходной точки разными методами: ГМ, МНК, НМ и ДМНК в оптимальном варианте, представляющие собой ломаные линии. Около каждой точки излома проставлен номер шага. Ход оптимизации по шагам отражен также в табл.1.1.

Таблица 1.1. Этапы оптимизации

Номер шага	Методы оптимизации								
	ГМ			МНК≡НМ			ДМНК		
	φ	x_1	x_2	φ	x_1	x_2	φ	x_1	x_2
0	909,00	-2,000	1,000	909,00	-2,000	1,000	909,00	-2,000	1,000
1	513,49	-0,781	1,304	778,25	-1,326	-1,019	52,49	-0,006	-0,717
2	11,41	-1,151	1,063	609,88	-0,525	-2,190	11,99	0,736	0,198
3	4,19	-0,013	1,121	355,87	0,555	-1,578	1,49	0,694	0,598
4	4,18	-1,057	1,096	26,660	1,228	2,024	0,29	0,852	0,674
5	4,19	-1,041	1,102	3,67	0,949	0,710	0,05	0,882	0,797
6	4,18	-1,046	1,097	0,39	1,017	1,098	0,01	0,931	0,858
7	4,17	-1,040	1,096	0,04	0,994	0,967	0,004	0,956	0,918
8	4,17	-1,043	1,091	0,01	1,001	1,011	0,002	0,975	0,947
9	4,16	-1,038	1,089	0,0003	0,999	0,996	0,00007	0,989	0,991
10	4,15	-1,040	1,083	-	-	-	-	-	-
100	3,54	-0,884	0,784	-	-	-	-	-	-
200	2,50	-0,596	0,354	-	-	-	-	-	-
500	0,06	0,749	0,558	-	-	-	-	-	-
4000	0,00007	0,991	0,985	-	-	-	-	-	-
Цена поиска (в пробах)	20000			56			50		

Из представленных результатов легко увидеть характерные особенности каждого метода. ГМ за каждый шаг продвигается на небольшую величину, в результате его траектория; есть зигзагообразная линия с мелким шагом; для приближения к минимуму с заданной точностью потребовалось около 4000 шагов, общая цена поиска для ГМ составила 20 000 проб.

МНК, который для данной системы идентичен НМ, оказался гораздо эффективнее, он потребовал девять шагов при цене 56 проб. По виду траектории, этого метода прослеживается, его тенденция к расходимости, обнаруживаемая в большом размахе зигзага, больших выбросах его за пределы долины. В случае более нелинейной системы эти выбросы сказались бы на длине каждого шага, т.е. на количестве шагов и цене поиска.

Наиболее эффективным для данного примера оказался ДМНК, траектория которого представляет собой криволинейную ломаную линию, идущую почти по долине. Количество шагов и цена поиска наименьшие – 9 и 50 соответственно. На траектории первого шага ясно видно, что в начальной части она касается траектории ГМ, а в конечной приближается к МНК.

На рисунке не показана полностью траектория СГМ, во избежание излишнего затемнения чертежа. Этот метод в данном примере оказался также неэффективен, как и ГМ, с тем отличием, что при приближении к минимуму его траектория имела тенденцию закручиваться вокруг точки минимума по спирали.

Следует сказать, что по опыту работы авторов с различными методами оптимизации на различных системах от простых, подобных приведенному примеру, до сложных, размерностью до 100×50, описанное поведение методов не является особенностью данного примера, но наоборот, отмеченные черты поведения типичны для подавляющего большинства оптимизируемых систем.

1.9. УСОВЕРШЕНСТВОВАНИЕ ДЕМПФИРОВАННОГО МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Из рассмотренного материала вытекает, что наиболее эффективным для оптимизации является демпфированный метод наименьших квадратов (ДМНК) с движением по кривой, траектория которого описывается выражением (1.43). Существенным недостатком этого метода является его сравнительная трудоемкость, необходимость при движении по траектории для определения вектора $\Delta X(\rho)$ каждой новой точки на траектории решать линейную систему порядка n, т.е. выполнять $\sim n^3$ действий. По трудоемкости такое количество действий сравнимо с пробой средней величины.

Ниже излагается предложенный авторами способ уменьшения трудоемкости метода, основанный на приведении матрицы M к диагональной форме.

В п.1.2 мы уже встречались с подобным приемом. Матрица M – симметрическая действительная неотрицательно определенная.

Из теории матриц следует [2, 7], что ее можно представить в виде произведения

$$M = U\Lambda U^T, \tag{1.59}$$

где U – ортонормированная действительная матрица собственных векторов; Λ – диагональная действительная неотрицательная матрица собственных значений. (Ортонормированная матрица отвечает условию $U^T U = UU^T = I$, где I – единичная матрица, или $U^{-1} = U^T$, т.е. матрица, обратная ортонормированной, совпадает с транспонированной).

Выражение (1.59) позволяет любую рациональную функцию матрицы $f(M)$ выразить через такую же функцию диагональной матрицы собственных значений $f(\Lambda)$

$$f(M) = Uf(\Lambda)U^T. \tag{1.60}$$

Поскольку любые функции диагональной матрицы находятся элементарно, полученные выражения могут быть использованы для простого вычисления матрицы $M_p^{-1} = (M + p^2 I)^{-1}$, необходимой при движении по траектории ДМНК,

$$M_p^{-1} = (M + p^2 I)^{-1} = U\Lambda_p^{-1}U^T, \tag{1.61}$$

где $\Lambda_p^{-1} = (\Lambda + p^2 I)^{-1} = \begin{pmatrix} \frac{1}{\lambda_1 + p^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_n + p^2} \end{pmatrix}$.

В справедливости (1.61) легко убедиться

$$(M + p^2 I)^{-1} = (U\Lambda U^T + p^2)^{-1} = (U\Lambda U^T + p^2 UIU^T)^{-1} =$$
$$= \left[U(\Lambda + p^2 I)U^T\right]^{-1} = (U^T)^{-1}(\Lambda + p^2 I)^{-1}U^{-1} = U\Lambda_p^{-1}U^T.$$

С использованием (1.61) траектория движения ДМНК может быть записана в виде

$$\Delta X(\rho) = -U\left[\Lambda + p^2(\rho)I\right]^{-1}U^T N. \tag{1.62}$$

Таким образом, движение по траектории требует для каждого значения ρ уже не решения линейной системы ($\sim n^3$ действий), а перемножения матриц $U\left[\Lambda + p^2(\rho)I\right]^{-1}U^T N$, на что расходуется $\sim n^2$ действий, т.е. в n раз меньше.

Для нахождения U и Λ можно воспользоваться методом вращений [3]. Так как этот метод итерационный, его трудоемкость зависит от заданной

точности. Применительно к целям оптимизаций авторами были получены экспериментальные данные, показывающие, что даже весьма низкая точность в определении значений U и Λ практически не влияет на сходимость метода, благодаря чему трудоемкость метода вращении оказалась не намного выше трудоемкости решения линейной системы (в 2–2.5 раза).

Сравнивая трудоемкость движения по траектории в старом и усовершенствованном ДМНК, получаем (C – количество действий): для старого $C \sim \ln^3$; для усовершенствованного $C \sim 2n^3 + \ln^3$; l – количество точек на траектории, обычно $l = 4 \div 5$. Таким образом, при $l > 2$ усовершенствованный метод требует меньше действий, чем обычный, причем его преимущество растет с увеличением l и n.

Вопрос о возможном дальнейшем сокращении трудоемкости метода за счет снижения требований к точности построения траектории без ухудшения сходимости представляет, по мнению авторов, существенный интерес и нуждается в специальном исследовании.

1.10. Список литературы.

1. Алгоритмы и программы случайного поиска [Сборник статей АН Латв. ССР]. Институт Электроники и вычислительной техники. Рига, «Зинатне», 1969. 372 с.
2. Беллман Р. Введение в теорию матриц. М., «Наука»б 1969. 368 с.
3. Березин И. С., Жидков Н. П. Методы вычислений. Т.2, М., Физматгиз, 1962. 640 с.
4. Гельфанд И. М., Цейтлин М. Л. Принцип нелокального поиска в системах автоматической оптимизации. ДАН СССР, вып. 137, 1969, № 2, с. 27–35.
5. Родионов С. А., Гутман Е. И. Основные вопросы построения стандартной программы «Оптимизатор». – «Известия вузов. Приборостроение», 1973, № 3, с. 71–75.
6. Уайлд Д. Дж. Методы поиска экстремума. Пер. с англ., М., «Наука», 1967. 267 с.
7. Форсайт Дж., Модер К. Численное решение систем линейных алгебраических уравнений. М., «Мир», 1969, 166 с.
8. Crockett I. B., Chernoff H. Gradient Methods of Maximization Pacific. J. Math. 5, 1955, p. 33.
9. Feder D. P. Automatic Optical, v. 2, N 12, 1963.
10. Lehman C. A. Designing Lens with a Computer Journal of the SMPTE 76, 1967, N 3, p. 188.
11. Kidger M. I., Wynne C. G. Experiments with Lens Optimization procedures. Optica Acta. 14, 1967, N 3, p. 279
12. Levenberg K. A Method for the Solution of certain non-linear problems in the last squares. Quarterly of Applied Mathematics v.2, 1944, N 2, p. 164.
13. Wynn C. G., Wormell I. H. Lens Design by Computer. Applied Optics 2, 1963, N 12, p. 1233.