МЕТОДЫ ОПТИМИЗАЦИИ ТЕХНОЛОГИЧЕСКИХ ЦИКЛОВ И ТЕПЛОВЫХ СХЕМ ТЭЦ

Грибунов Г.А., Куликов И.С.

В теплоэнергетике приходится соприкасаться с большим разнообразием подлежащих оптимизации задач (например, выбор мощности и типа нового оборудования, рабочих и эксплуатационных параметров технологических циклов и процессов, тепловой схемы установки, параметров схемы теплоснабжения и др.). В зависимости от сложности рассматриваемых задач, цели исследования и уровня изученности исходных данных могут применяться частные либо комплексные методы оптимизации. Поскольку в энергетических установках большинство параметров и элементов взаимосвязаны, оптимальное решение по оптимизируемому объему в целом может быть получено лишь на основе комплексной оптимизации. В оптимизационных технико-экономических расчетах, как правило, используются два метода - вариантный и аналитический [1].

Вариантный метод прост, нагляден, не требует сложных математических зависимостей, однако без предварительной оптимизации каждого из рассматриваемых вариантов может приводить к искаженным результатам. Он применяется в основном при проектировании. В качестве критерия оптимума здесь используются приведенные расчетные затраты ***З*** либо их переменная часть Δ***З***.

Аналитические методы базируются на применении математических зависимостей, устанавливающих непосредственную взаимосвязь критериев оптимума и оптимизируемых параметров. Установление функциональной зависимости критерия оптимума от искомого параметра при аналитической оптимизации является достаточно сложной физической и математической задачей. Однако широкие возможности использования в исследованиях ЭВМ и разработка методов математического моделирования и анализа в энергетике сделали аналитические методы основными современными методами оптимизации. В аналитических методах оптимизации критериями оптимальности могут быть, например, коэффициент термодинамической эффективности $\eta_{Т.Э.}$ (термодинамическая оптимизация), приведенные затраты ***З*** или их переменная часть Δ***З*** (технико-экономическая оптимизация). Оптимальное значение искомого параметра x_i соответствует экстремуму функции:

$$\left(\frac{\partial \eta_{Т.Э.}}{\partial x_i}\right)_j = 0; \qquad \left(\frac{\partial З}{\partial x_i^2}\right)_j = 0$$

при условии

$$\left(\frac{\partial^2 \eta_{Т.Э.}}{\partial x_i^2}\right)_j = 0; \qquad \left(\frac{\partial^2 З}{\partial x_i^2}\right)_j = 0$$

где j – фиксированные параметры, независимые от x_i.

Если осуществляется комплексная оптимизация нескольких параметров $(x_1, x_2, ..., x_n)$, то критерий оптимальности дифференцируется отдельно по каждому из этих параметров:

$$\left(\frac{\partial З}{\partial x_1}\right) x_2, x_3, ..., x_n = 0;$$

$$\left(\frac{\partial З}{\partial x_2}\right) x_1, x_3, ..., x_n = 0; \tag{1}$$

$$\left(\frac{\partial З}{\partial x_n}\right)x_1, x_2, ..., x_{n-1} = 0.$$

При наличии исходной взаимосвязи, влияющей на результаты оптимизации между оптимизируемыми параметрами, в систему (1) вводится уравнение связи

$$\Phi = f\left(x_1, x_2, x_3, ..., x_n\right).$$

В этом случае искомые параметры могут быть найдены, например, с помощью метода множителей Лагранжа:

$$\frac{\partial З}{\partial x_1} + \lambda\frac{\partial \Phi}{\partial x_1} = 0;$$

$$\frac{\partial З}{\partial x_2} + \lambda\frac{\partial \Phi}{\partial x_2} = 0;$$

$$\frac{\partial З}{\partial x_n} + \lambda\frac{\partial \Phi}{\partial x_n} = 0,$$

где λ – неопределенный множитель Лагранжа.

В практике оптимизации следует пользоваться следующими методическими рекомендациями. Если расчетные экономические и термодинамические зависимости оказываются слишком сложными, то можно находить относительное изменение приведенных затрат рассматриваемых вариантов $\partial З$ по отношению к какому-то исходному, так называемому базовому варианту (либо базовой точке). В качестве базового можно принять любой технически оправданный вариант, для которого известны технические характеристики и параметры, зависящие от искомого параметра, и рассматриваются все составляющие расчетных затрат при каком-то значении искомого параметра $x = x_0$, подлежащего в дальнейшем оптимизации. При оптимизации по базовому варианту существенно упрощаются расчетные зависимости, уменьшаются объемы вычислений и повышается точность расчетов. Этим методом удобно пользоваться при оптимизации эксплуатационных параметров энергетических установок [1].

При оптимизации параметров, характеристик и режимов работы оборудования необходимо принять возможные технические ограничения, от которых существенно зависят результаты оптимизации. В ряде случаев технические ограничения не позволяют применять комплексную оптимизацию. Под техническими ограничениями понимают такие, при которых ограничено изменение самих исследуемых параметров либо ограничены промежуточные величины, влияющие на выбор оптимизируемых параметров. В этом случае в аналитические методы оптимизации вводятся дополнительно условия, учитывающие технические ограничения.

В теплоэнергетике при оптимизации нужно учитывать технические ограничения в случае выбора числа ступеней подогрева питательной и сетевой воды, числа выхлопов турбины, промежуточного перегрева пара, числа трубопроводов (при оптимизации скорости движения рабочего тела и теплоносителя), характеристик тягодутьевых и водоохладительных устройств, числа типоразмеров унифицируемого оборудования (котлоагрегаты, подогреватели, насосы, ЦНД турбин и др.). Характеристики перечисленных объектов изменяются дискретно от одного целого числа к другому. Это значит, что при исследовании непрерывной функции на экстремум физический и экономический оптимумы будут иметь лишь те значения искомого параметра, которые равны целому числу.

Большое распространение получил в оптимизационных исследованиях **метод математического моделирования**. С помощью этого метода, основанного на использовании комплексов алгоритмов и программ, решаются любые технико-экономические оптимизационные задачи. Применение ЭВМ позволяет решать систему из множества комплексных уравнений, описывающих с достаточной точностью зависимость физических параметров теплоносителей и рабочих тел, термодинамические, расходные, конструктивные, режимные и прочие характеристики объектов оптимизации, а также их взаимосвязь.

Математическое моделирование подразумевает неразрывную связь между экономико-математической моделью и методом решения ее. Первым этапом математического программирования является построение математической модели исследуемого объекта. Математическая модель – это система математических уравнений, описывающих наиболее характерные черты функционирования и развития моделируемого объекта. Математическая модель должна удовлетворять двум противоречивым требованиям: с одной стороны, она должна как можно точнее описывать экономическую систему; с другой - должна быть достаточно простой, чтобы ее можно было описать с помощью математических уравнений и чтобы получаемая система уравнений была доступна для использования имеющимися математическими методами и ЭВМ. Поэтому математическая модель описывает экономическую систему приближенно, выделяя одни ее свойства и пренебрегая другими. В общем случае для одного и того же объекта можно построить большое число различных математических моделей, так как построение модели зависит не только от природы и качества объекта, но и от цели (критерия эффективности), которой она должна соответствовать. В качестве критерия эффективности наибольшее распространение получил минимум расчетных затрат, хотя в ряде случаев могут быть использованы и другие критерии, например надежность, максимум прибыли, минимум себестоимости, минимум расхода топлива. Математическая формулировка зависимости критерия эффективности от параметров модели называется целевой функцией (целевым функционалом) модели. Часто в экономических задачах область изменения целевой функции или переменных стеснена рядом условий, имеющих вид равенств или неравенств. Такие условия называются ограничивающими, или ограничениями модели. Таким образом, математическая модель предполагает наличие целевой функции и системы ограничивающих условий.

Под оптимальностью понимается получение такого решения, которое экстремизирует (максимизирует или минимизирует) целевую функцию при выполнении заданных ограничений.

По характеру использования математического аппарата и с точки зрения принципов формирования различают следующие математические модели:

– ***линейные***, в которых все математические зависимости, характеризующие объект, описываются линейными функциями, и ***нелинейные***, в которых все или часть зависимостей описываются нелинейными функциями;

– ***статические***, в которых рассматривается один фиксированный отрезок времени, и ***динамические***, которые описывают развитие системы во времени;

– ***детерминированные***, не учитывающие случайного характера реальных процессов, и ***вероятностные*** *(стохастические, статистические)*, которые учитывают случайный характер протекающих процессов;

– ***непрерывные***, в которых параметры модели могут принимать любые численные значения, и ***дискретные*** *(целочисленные)*, в которых переменные могут быть лишь целыми числами;

– ***оптимизационные***, в которых оптимальный вариант определяют исходя из экстремизации целевой функции, и ***оценочные*** *(расчетные)*, основанные на анализе нескольких возможных решений с целью выбора наилучшего;

– ***глобальные***, непосредственно дающие экстремум целевой функции, и ***блочные***, экстремум целевой функции которых достигается лишь при взаимодействии блоков модели [1].

Наибольшее развитие получили методы линейного программирования. Уровень разработки методов линейного программирования позволяет решать большинство практических задач. Однако применение линейного программирования требует линеаризации реальных зависимостей, имеющих, как правило, нелинейный характер.

Постановка задачи линейного программирования в общем виде записывается системой: минимизировать

$$f\left(x_1,...,x_n\right)=\sum_{i=1}^{n}c_i x_i$$

при ограничениях

$$\sum_{i=1}^{n}a_{ij}x-b_j\geq 0,\quad j=\overline{1,m}; \tag{2}$$

$$x_i\geq 0,\qquad i=\overline{1,n}, \tag{3}$$

где a,b,c – константы; x_i – искомые переменные.

Одним из универсальных методов решения задач линейного программирования является симплексный метод [2], называемый часто методом последовательного улучшения плана. Для решения задачи с помощью симплекс-метода необходимо прежде всего привести систему ограничений (2) к канонической форме, т. е.

$$\sum_{i=1}^{n}a_{ij}x_i+y_j=b_j. \tag{4}$$

Затем, решая полученную систему (4) относительно n переменных, найдем допустимое базисное решение:

$$y_j=b_j-\sum_{i=1}^{n}a_{ij}x_i.$$

Алгоритм симплекс-метода состоит из двух этапов: на первом находится допустимое базисное решение; на втором осуществляется переход от начального решения к другому, более близкому к оптимальному, затем к следующему, и так до тех пор, пока задача не будет решена. С геометрической точки зрения система ограничений представляет область выпуклого многогранника допустимых решений. Базисное решение является вершиной этого многогранника. Переход от одного решения к другому можно толковать как перемещение по ребрам от одной вершины многогранника к другой по направлению к той вершине, которая соответствует наилучшему решению.

Кроме симплекс-метода, существуют и другие вычислительные методы линейного программирования.

Обобщением задач линейного программирования является параметрическое программирование. Особенность задачи параметрического программирования состоит в том, что исходные данные являются не постоянными величинами, а функциями, зависящими от некоторых параметров. Рассмотрим случай, когда целевая функция линейно зависит от параметра p. Математическая формулировка задачи для параметра p, принадлежащего отрезку $kb, p\in\left(k,b\right)$, где k,b – произвольные действительные числа, выглядит так: найти для каждого p на отрезке $\left(k,b\right)$ такой вектор $x^*=\left(x_1^*,...,x_m^*\right)$, который максимизировал бы целевую функцию

$$\sum_{i=1}^{m}\left(c_i+d_i p\right)\,x_i$$

при ограничениях

$$\sum_{i=1}^{m}a_{ij}x_i\leq a_j;\quad j=\overline{1,n};\quad x_i\geq 0;\quad i=\overline{1,m}.$$

Решение задачи состоит из двух этапов. На первом этапе фиксируем значение параметра p, например, задаем $p=k$. Таким образом, получаем обычную задачу линейного программирования. Решая ее тем или иным методом, определяем значение x_i, максимизирующее целевую функцию при заданных ограничениях. На втором этапе определяется интервал изменений параметра p_i, для которого максимум целевой функции достигается в одной и той же вершине многогранника допустимых решений. Найденный интервал исключается из отрезка (k,b). Для оставшейся части отрезка вновь повторяется процедура первого этапа задачи. Решение будет повторяться до тех пор, пока отрезок (k,b) не будет полностью разбит на интервалы.

Постановка задачи линейного программирования требует предварительной линеаризации реальных зависимостей, которые являются, как правило, нелинейными. Если линеаризация приводит к значительной погрешности, то для решения поставленной задачи необходимо использовать методы нелинейного программирования. Общая задача нелинейного программирования заключается в отыскании экстремума целевой функции при заданных ограничениях, причем либо оптимизируется функция, либо система ограничений, либо та и другая вместе нелинейные.

Обычно в качестве отдельной группы выделяют методы, разработанные в классической математике: метод поиска оптимума путем решения системы нелинейных уравнений, полученных при приравнивании нулю частных производных исследуемой функции по оптимизируемым параметрам, и метод неопределенных множителей Лагранжа. Эти методы позволяют решать задачи поиска оптимума нелинейной функции многих переменных только при отсутствии ограничений на оптимизируемые параметры или при ограничениях в виде равенств. Поэтому указанные методы нельзя относить к методам нелинейного математического программирования.

Очень часто собственно методы нелинейного математического программирования, т. е. методы определения экстремума нелинейной функции при наличии ограничений на оптимизируемые параметры в виде неравенств, делят по признаку организации процесса поиска на методы слепого поиска и методы направленного поиска. К ***методам слепого поиска*** относятся метод сплошного перебора вариантов (метод прямого упорядочения вариантов по критерию эффективности) и метод статистических испытаний (метод Монте-Карло). К ***методам направленного поиска*** относятся градиентный метод, метод наискорейшего спуска, метод покоординатного спуска и др. [3].

Согласно другой классификации все методы нелинейного программирования можно разделить на методы локального поиска и методы нелокального (глобального) поиска. В процессе решения задачи одним из ***локальных методов*** значения оптимизируемых параметров непрерывно меняются в направлении минимизации (или максимизации) рассматриваемой функции. Тем самым эти методы гарантируют нахождение только локального оптимума. К группе локальных методов относятся методы: градиентный, наискорейшего спуска, покоординатного спуска и др. Для методов ***глобального поиска*** характерно введение дискретности в процессе изменения оптимизируемых параметров, что способствует рассмотрению большей области изменения исследуемой функции и выявлению абсолютного оптимума среди локальных. К этой группе методов относятся метод случайного поиска, метод динамического программирования, а также сочетания для совместного использования ряда других методов.

Классические методы решения экстремальных задач. К числу классических математических методов определения экстремума функции многих переменных относятся:

1) метод поиска оптимума путем решения системы нелинейных уравнений, полученных при приравнивании нулю частных производных минимизируемой функции по оптимизируемым параметрам;

2) метод неопределенных множителей Лагранжа.

В математическом плане оба метода достаточно хорошо известны, поэтому основное внимание при их анализе следует обратить на вопрос применимости этих методов к решению задач оптимизации теплоэнергетических установок.

Рассмотрим первый из указанных методов. Пусть имеем функцию цели $З(x_1, x_2, ..., x_n)$. Расположение экстремума внутри области можно определить по следующему правилу: непрерывная функция $З(x_1, x_2, ..., x_n)$ от n независимых переменных $x_1, x_2, ..., x_n$ достигает максимума или минимума внутри области только при таких значениях переменных x_i, для которых n частных производных $\frac{\partial З}{\partial x_1}, \frac{\partial З}{\partial x_2}, ..., \frac{\partial З}{\partial x_n}$ одновременно обращаются в нуль. Согласно этому правилу необходимым условием определения экстремальных точек является решение системы n алгебраических или трансцендентных уравнений, получаемых приравниванием нулю частных производных:

$$\frac{\partial З}{\partial x_1} = 0; \quad \frac{\partial З}{\partial x_2} = 0; \quad ...; \quad \frac{\partial З}{\partial x_n} = 0. \tag{5}$$

На первый взгляд кажется, что использование этого метода позволяет достаточно просто решать задачу определения оптимума нелинейной функции многих переменных. Однако это не так. Существует ряд трудностей при его реализации и ограничений по сфере его применения. Во-первых, при большом числе оптимизируемых параметров рассматриваемый метод становится весьма сложным в части решения системы уравнений (5). Задача решения системы уравнений (5) только в простейших случаях оказывается легкой. В практических задачах оптимизации теплоэнергетических установок число переменных x_i, как правило, велико. Во-вторых, условие определения экстремума, выраженное зависимостью (5), является необходимым, но недостаточным для решения задачи. В самом деле, выражение (5) определяет положение стационарных точек внутри области, среди которых кроме экстремальных могут быть особые точки типа «седла». Учет достаточных условий нахождения экстремумов функции многих переменных является весьма сложным как в алгоритмическом, так и в вычислительном плане. В-третьих, рассматриваемый метод дает возможность найти экстремум только в том случае, если он лежит внутри, а не на границе области возможных значений аргументов. Между тем многие параметры и характеристики теплоэнергетических установок имеют свои оптимальные значения именно на границах допустимой области их изменения. Следовательно, требуется дополнительный анализ значений минимизируемой функции $З(x_1, x_2, ..., x_n)$. на границах допустимой области изменения параметров $x_1, x_2, ..., x_n$. Наконец, четвертый недостаток рассматриваемого метода состоит в ограниченности его применения классом задач, в которых оптимизируемые параметры, определяющие значение минимума или максимума функции, независимы, т. е. $x_1, x_2, ..., x_n$ независимы друг от друга. В то же время для теплоэнергетических установок, отдельных агрегатов и элементов оборудования характерно наличие балансовых уравнений, связывающих между собой значительную часть параметров [4].

Все указанные недостатки приводят к выводу о том, что использование классического метода определения экстремумов функции многих переменных для решения задач оптимизации параметров теплоэнергетических установок или отдельных элементов является неэффективным, поскольку:

1) оно сводит первоначально поставленную задачу отыскания экстремума к таким вторичным задачам, которые оказываются не проще исходной, а зачастую и сложнее;

2) при этом возникает необходимость в значительном изменении условий постановки энергетической задачи, искажающем ее сущность.

Определение экстремальных точек функции многих переменных для весьма важного случая наличия дополнительных связей между оптимизируемыми параметрами может быть осуществлено с использованием классического математического метода множителей Лагранжа [1]. Пусть имеем функцию цели $З(x_1, x_2, ..., x_n)$, экстремум которой требуется найти, причем имеют место дополнительные условия

$$\varphi_j(x_1, x_2, ..., x_n) = 0, \tag{6}$$

где $j = 1, 2, ..., m$; $m < n$.

Введя m дополнительных множителей, $\lambda_1, \lambda_2, ..., \lambda_m$ и обозначив $X = (x_1, x_2, ..., x_n)$ и $\Lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$, получаем новую функцию:

$$L(X, \Lambda) = З(X) + \sum_{j=1}^{m} \lambda_j \varphi_j(X), \tag{7}$$

где λ_j – множитель Лагранжа.

Необходимые условия экстремума состоят в равенстве нулю всех первых частных производных от L. В результате получается $(n+m)$ уравнений с $(n+m)$ неизвестными X и Λ. Решение этих уравнений относительно переменных X и Λ позволяет определить положение стационарной точки. Таким образом, использование вспомогательной функции $L(X, \Lambda)$ и вспомогательных множителей Λ позволяет заменить задачу с дополнительными условиями вида (6) задачей без дополнительных условий.

Недостатком метода множителей Лагранжа является введение m дополнительных переменных, которые должны быть исключены с помощью m дополнительных уравнений. Если учесть, что при решении задачи комплексной оптимизации параметров теплоэнергетических установок число уравнений связи между оптимизируемыми параметрами достигает 100–200, то станет очевидной важность этого недостатка. Кроме отмеченного для метода множителей Лагранжа сохраняют свою силу и другие недостатки и трудности использования, указанные выше применительно к первому из рассмотренных методов.

Методы слепого поиска. Сущность решения задач на определение экстремума функции многих переменных $З(x_1, x_2, ..., x_n)$ с помощью методов слепого поиска заключается в организации просмотра (в определенном порядке или случайным образом) допустимой области изменения оптимизируемых параметров X и сравнении соответствующих значений функции $З$. При этом информация о функции $З$, полученная в результате какого-либо варианта расчета, используется при последующем расчете лишь в ограниченном объеме, а само определение экстремального значения функции $З$ не сопровождается последовательным улучшением промежуточных результатов.

К числу методов слепого поиска принадлежит ***метод прямого упорядочения вариантов по критерию эффективности*** (метод пространственной сетки). Суть его состоит в следующем. Для каждого независимого оптимизируемого параметра x_i технически допустимая зона определения делится на равные отрезки. Значения параметров на концах полученных отрезков образуют новую, уже дискретную область определения этих параметров. Число отрезков выбирается по допустимому количеству точек дискретной области определения функции $З$:

$$N = m_1 m_2, ..., m_i, ..., m_n, \tag{8}$$

где m_i – количество точек дискретной области определения i-го параметра; n - число независимых оптимизируемых параметров.

Наибольшее значение $N_{макс}$ определяется допустимым по техническим или экономическим соображениям временем работы ЭВМ:

$$t_{доп} = N_{макс}\tau, \tag{9}$$

где τ – время расчета одного варианта.

В процессе последовательного расчета вариантов очередное значение функции ***З*** сравнивается с минимальным из ранее рассмотренных и в результате выбирается экстремальное значение (зона значений) целевой функции ***З***. Варианты, не удовлетворившие тем или иным ограничениям, поставленным в условиях задачи, из сопоставления исключаются. При решении задач выпуклого нелинейного программирования методом последовательного сравнения вариантов способ деления допустимой зоны определения каждого независимого оптимизируемого параметра на отрезки равной длины не является наилучшим. Целесообразнее проводить поиск экстремума при переменной длине отрезка, уменьшая его по мере приближения к зоне оптимума. Сопоставление ряда способов выбора размера отрезка показывает, что для задач этого класса оптимальным является способ деления, в основу которого положена последовательность чисел Фибоначчи.

К преимуществам метода прямого упорядочения вариантов по критерию эффективности следует отнести простоту алгоритма и программы оптимизации, малый объем необходимой машинной памяти и возможность нахождения абсолютного оптимума. Главным недостатком метода является большое время работы ЭВМ, так как приходится рассчитывать все возможные варианты сочетаний значений оптимизируемых параметров. Этот недостаток вытекает из самой сущности рассматриваемого метода, при котором в процессе поиска экстремального значения целевой функции ***З*** результаты расчета предыдущих вариантов используются в очень малой степени. Кроме того, рассматриваемый метод позволяет определить лишь приближенное положение точки оптимума, соответствующее значению функции цели в узлах пространственной сетки.

Применение рассматриваемого метода целесообразно для оптимизации параметров элементов и групп элементов оборудования теплоэнергетических установок, имеющих относительно небольшое число независимых параметров и варьируемых внешних факторов.

Следующим методом слепого поиска, который может быть применен в процессе оптимизации параметров теплоэнергетических установок и их отдельных элементов для решения нелинейных, экстремальных, многофакторных задач, является ***метод статистических испытаний*** **(метод Монте-Карло)**. Сущность этого метода заключается в том, что решение аналитической задачи заменяется моделированием некоторого случайного процесса. Его вероятностная характеристика, например вероятность определенного события или математического ожидания некоторой величины, имеет тесную связь с возможным решением исходной аналитической задачи. При использовании указанного метода необходимо большое число раз моделировать соответствующий случайный процесс и определять путем статистической обработки значение искомой характеристики – вероятности или математического ожидания. Поэтому метод статистических испытаний требует выполнения огромной вычислительной работы.

Метод статистических испытаний характеризуется простотой алгоритма и программы решения задачи. Ему свойственны все преимущества, присущие методу прямого упорядочения вариантов по критерию эффективности. Вместе с тем при использовании метода статистических испытаний количество рассчитываемых вариантов, а, следовательно, и время счета на ЭВМ зависят от требуемой вероятности решения задачи с погрешностью, не превышающей определенное значение. Для тех задач, где допустимо некоторое снижение вероятности получения решения с заданной точностью, число необходимых случайных испытаний может быть уменьшено.

Анализ возможностей использования двух методов слепого поиска для решения многофакторных экстремальных задач показал, с одной стороны, ряд их положительных свойств, а с другой – ограниченность их применения кругом задач с небольшим числом оптимизируемых параметров. Второй весьма важной областью применения методов слепого поиска является их использование в алгоритмах, сочетающих в себе ряд методов, в частности для определения абсолютного оптимума в многоэкстремальных задачах и для оптимизации дискретно изменяющихся параметров.

Методы направленного поиска. Для оптимизации теплоэнергетических установок и их отдельных элементов с большим числом оптимизируемых параметров и варьируемых факторов могут быть применены методы направленного поиска оптимума: градиентные, наискорейшего спуска, покоординатного спуска и др. Характерной чертой этих методов является использование в процессе решения задачи результатов каждого данного шага (иногда также и предыдущих шагов) поиска оптимальной точки для определения направления изменения оптимизируемых параметров на каждом следующем шаге. При этом значение минимизируемой функции систематически уменьшается. Тем самым вместо рассмотрения большого количества вариантов происходит направленный анализ относительно малого числа вариантов, что обеспечивает приемлемое время счета на современных ЭВМ среднего класса для задач, в которых число оптимизируемых параметров исчисляется десятками.

Рассмотрим *градиентный метод* для простейшего случая определения экстремума функции многих переменных $З(x_1, x_2, ..., x_n)$ при отсутствии каких-либо ограничений. Процесс оптимизации по методу градиента заключается в определении направления наискорейшего изменения функции ***З*** и некотором перемещений по этому направлению в прямую или обратную сторону. Направление наискорейшего изменения функции определяется направлением вектор-градиента оптимизируемой функции, которое всегда совпадает с направлением возрастания функции. Компонентами градиента $\frac{\partial З}{\partial X^0}$ в какой-либо точке рассматриваемой области заданной параметрами $(x_1^0, x_2^0, ..., x_n^0)$, являются частные производные функции $\frac{\partial З}{\partial x_1^0}, \frac{\partial З}{\partial x_2^0}, ..., \frac{\partial З}{\partial x_n^0}$.

Отметим, что градиент $\frac{\partial З}{\partial X^0}$ всегда перпендикулярен к поверхности равных значений функции ***З*** в рассматриваемой точке.

Каждый этап процесса градиентного спуска имеет две составляющие: определение направления наискорейшего спуска и осуществление шага по направлению спуска.

Наименьшее число шагов при градиентном спуске обеспечивает разновидность градиентного метода, называемая иногда ***методом наискорейшего спуска***. Суть этой модификации градиентного метода в следующем. После определения градиента функции $З(X)$ производится движение по направлению антиградиента до точки, в которой достигается минимальное значение функции $З(X)$ на данном направлении. В найденной точке снова определяется градиент и движение совершается по прямой, соответствующей направлению нового антиградиента, и т. д. до нахождения экстремума функции $З(X)$.

Для определения шага при использовании метода наискорейшего спуска могут быть применены два способа. Первый из них исходит из положения, что на каждом прямом участке линии антиградиента функция $З(X)$ может рассматриваться как параметрическая функция одного параметра t, т. е. $З = З\left[X(t)\right]$. Тогда для вычисления шага, конец которого совпадает с минимумом функции $З\left[X(t)\right]$ на этой линии, достаточно найти минимум функции $З\left[X(t)\right]$ по параметру t.

Второй способ определения шага в методе наискорейшего спуска базируется на интерполяции (экстраполяции) изменения функции $З(X)$ вдоль направления антиградиента.

Рассматриваемый способ определения шага наискорейшего спуска имеет определенную погрешность, вызываемую неточностью описания действительной зависимости изменения функции $З(X)$ по направлению антиградиента интерполирующим полиномом. Значение погрешности определения dX может быть снижено, если применить интерполяцию полиномом третьей или более высокой степени, но при этом потребуется вычислить значения $З(X)$ для большого числа точек на направлении антиградиента.

Анализ приведенных способов выбора шага в градиентном методе спуска к точке минимума не позволяет сделать однозначного заключения о безусловных преимуществах какого-либо одного из них. Причины этого достаточно очевидны. С одной стороны, от выбранного способа определения шага зависят сходимость вычислительного процесса, выражающаяся через число шагов, необходимых для достижения точки оптимума, и соответственно время счета на ЭВМ. С этой точки зрения более целесообразными являются два последних из рассмотренных способов, обеспечивающие решение задачи оптимизации за минимальное число шагов. Но, с, другой стороны, эти последние способы определения шага весьма сложны и могут потребовать значительного времени для расчета на ЭВМ собственно шага. Поэтому выбор способа определения шага должен производиться в каждом конкретном случае решения той или иной задачи с учетом инженерной специфики объекта оптимизации, объема задачи, требований к точности решения, характеристик используемой ЭВМ и других факторов.

Следующим достаточно эффективным методом направленного поиска является *метод покоординатного спуска* (метод Гаусса - Зейделя). Суть этого метода заключается в минимизации многопараметрической функции $З = З\left(x_1^0, x_2^0, \ldots, x_n^0\right)$, где индекс 0 обозначает принадлежность параметра к исходной точке спуска, сначала по одному параметру x_1, затем по второму x_2 и т. д. до последнего параметра x_n. На первом этапе решения задачи фиксируются значения всех параметров, кроме первого, и определяется оптимальное значение этого параметра, т. е. ищется минимум функции $З = З\left(x_1^0, x_{2\phi}^0, \ldots, x_{n\phi}^0\right)$. Найденное оптимальное значение первого параметра обозначим x_1^1. Далее ищется минимум функции $З\left(x_{1\phi}^1, x_2, x_{3\phi}^0, \ldots, x_{n\phi}^0\right)$ при изменении только второго параметра x_2. При этом первый параметр x_1 фиксируется при найденном выше оптимальном значении, т. е. $x_1 = x_{1\phi}^1$.

Цикл оптимизации заканчивается после определения минимума функции $З = З\left(x_{1\phi}^1, x_{2\phi}^1, \ldots, x_{(n-1)\phi}^1, x_n\right)$ при изменении параметра x_n, что соответствует установлению его оптимального значения. Один цикл поиска при использовании метода покоординатного спуска, т. е. однократная раздельная оптимизация значений всех параметров X, как правило, не позволяет найти состояние, соответствующее минимуму функции $З(X)$. Поэтому необходимо повторение указанного цикла.

Способы выбора длины шага при покоординатном спуске совпадают со способами, рассмотренными применительно к градиентному методу. Последовательность, в которой выбираются координатные оси, может быть различной. Обычно они берутся в фиксированном циклическом порядке (чаще всего просто поочередно). Иногда вначале определяется, какие из переменных $x_1, x_2, ..., x_n$ оказывают наибольшее влияние на изменение функции цели $З$, в соответствии с чем и строится последовательность спуска по отдельным параметрам. Такое предварительное исследование и расстановка параметров в порядке их значимости безусловно повышают эффективность метода, улучшая его сходимость. Однако указанное предварительное исследование в свою очередь требует определенного усложнения алгоритма оптимизации и дополнительного времени счета на ЭВМ.

Имеющийся опыт применения метода покоординатного спуска показывает, что по условию сходимости он при малом числе переменных может дать лучшие результаты, чем градиентный метод. Однако при решении задач с большим числом переменных и сложной системой ограничений метод покоординатного спуска существенно уступает градиентному методу.

При большом числе оптимизируемых параметров X иногда применяют модифицированный метод покоординатного спуска. В этом случае полное число параметров X разбивается на некоторое число групп. Вначале одним из рассмотренных методов оптимизируются параметры первой группы при фиксированных значениях параметров остальных групп. Далее осуществляется поиск оптимума по параметрам второй группы и т. д. Таким образом, отличие этого метода от простого покоординатного спуска состоит в том, что при простом покоординатном спуске поиск экстремума идет поочередно по каждой переменной, а при модифицированном методе поочередно по группам переменных.

Следует остановиться еще на одной модификации метода покоординатного спуска – методе грубого поиска минимума, также нашедшем применение для решения задач большой размерности. Согласно этому методу первоначально производится спуск по всем переменным поочередно. Далее выбирается несколько координат, по которым спуск наиболее эффективен. Последующий спуск производится только по этим координатам (все остальные не меняются). Время от времени выбор этих ведущих координат повторяется заново.

Каждый из методов направленного поиска имеет упомянутые нами слабые и сильные стороны. Вместе с тем им свойственны общие для всех методов преимущества и недостатки. Основное их преимущество заключается в направленности поиска оптимума, что позволяет решать задачи с большим числом оптимизируемых параметров на ЭВМ среднего класса за приемлемое время. Именно это их достоинство обусловило широкое использование методов направленного поиска при решении экстремальных многофакторных задач. Среди недостатков методов направленного поиска следует выделить один основной – возможность нахождения только локального оптимума или особой точки типа седловой.

Метод случайного поиска. Общая их идея заключается в следующем. В процессе минимизации в окрестностях точки X^i определяется одно или несколько значений функции $З(X)$, для чего делаются соответствующие пробные шаги. Затем на основании полученных значений функции $З(X)$ вычисляется новая точка X^{i+1}, т. е. делается рабочий шаг. Далее процесс повторяется. При этом направления изменения компонент вектора X задаются случайными, причем все направления равновероятны, а движение к экстремуму будет осуществляться только в том случае, когда результат данного случайного движения приводит к уменьшению функции $З(X)$.

Следует заметить, что в инженерных задачах оптимизации параметров теплоэнергетических установок, при постановке которых широко используются предшествующий опыт и знания о близких по типу установках, практически не бывает случаев, когда исходная точка отстоит от точки оптимума на очень большие расстояния.

Преимущества градиентного метода оптимизации по сравнению с методом случайного поиска возрастают в случае организации процесса спуска с переменным рабочим шагом. Для этого случая в процессе случайного поиска среднее приращение функции $З(X)$ на один расчет $\frac{2n}{n+1}$ раз меньше, чем при градиентном методе, где n – число оптимизируемых параметров X. Принцип случайного поиска обладает важными преимуществами: во-первых, алгоритмы, его реализующие, менее чувствительны, чем детерминированные методы, к наличию неглубоких локальных минимумов и, во-вторых, некоторые алгоритмы случайного поиска позволяют определять точку абсолютного минимума.

Методы учета ограничений. До сих пор рассматривалось применение методов направленного и случайного поиска для простейшего случая оптимизации нелинейной функции $З(X)$ при отсутствии каких-либо ограничений. Более общим случаем является задача минимизации функции многих переменных при наличии ограничений в виде равенств. Эта задача может быть сформулирована следующим образом.

Необходимо минимизировать нелинейную функцию цели

$$З = З\left[X, Y(X)\right] \tag{10}$$

при наличии нелинейных ограничений в виде равенств

$$\Phi\left[X, Y(X)\right] = 0. \tag{11}$$

Наиболее общим случаем задачи минимизации нелинейной функции $З = З\left[X, Y(X)\right]$ является случай, когда кроме ограничений в виде системы нелинейных равенств $\Phi\left[X, Y(X)\right] = 0$ и ограничений на совокупность независимых параметров X имеют место ограничения на зависимые параметры $Y^* \le Y\left[X, Y(X)\right] \le Y^{**}$ и некоторые дополнительные функции $F: F^* \le F\left[X, Y(X)\right] \le F^{**}$. Здесь в качестве дополнительных функций F выступает совокупность нелинейных технологических характеристик, имеющих ограничения. Принципиальное отличие ограничений на Y и F от ограничений на параметры X состоит в том, что параметры Y и характеристики F являются нелинейными функциями параметров X и Y. Именно это обстоятельство обусловливает особую сложность решения задачи в такой общей постановке. Процесс градиентного спуска должен быть организован таким образом, чтобы при выходе из допустимой области дальнейшее движение происходило вдоль нелинейных границ и одновременно в направлении снижения значения функции $З$. В настоящее время известен ряд способов решения таких задач, некоторые из них будут рассмотрены ниже.

Способ зигзагообразного движения вдоль границ для простейшего случая двух переменных (x_1 и x_2) и одного ограничения $\left(f_1^*\right)$ иллюстрируется рис. 1. Пусть при движении по направлению антиградиента из точки, расположенной в допустимой области изменения параметров (точка A_2 на рис. 1), сделан шаг, при котором нарушаются некоторые из ограничений по F, т. е. имеет место нарушение границы. Для возврата в допустимую область организуется движение в направлении вектора S, определяемого как сумма градиентов тех функций f_ρ, которые в точке A_3 вышли за ограничения:

$$S = \sum_{\rho=1}^{N} \frac{\partial f_\rho}{\partial X}. \tag{12}$$

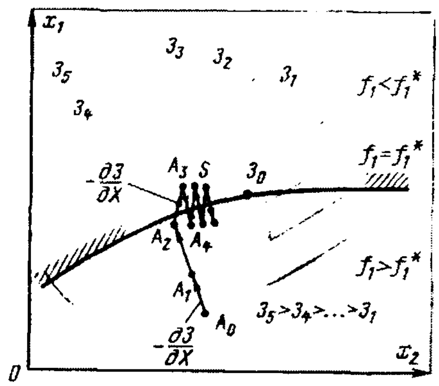

Рис.1. Зигзагообразное движение вдоль границы допустимой области

Это движение сменяется движением по направлению антиградиента, как только текущая точка поиска снова окажется в допустимой области и т. д. Из рис. 1 хорошо видно, что по мере приближения к цели (точка $З_0$) снижение минимизируемой функции $З$ постепенно замедляется. Медленная сходимость вычислительного процесса является основным недостатком рассмотренного метода. Аналогичным образом изложенный способ может быть использован для движения вдоль границ ограничений по Y и при одновременном учете ограничений по F и Y.

Определенные возможности для движения по границам допустимой области представляет метод проектирования градиента. Рассмотрим основы этого метода на простейшем примере минимизации функции $З(X)$ при наличии одного ограничения в виде неравенства на технологическую характеристику f_1 (рис. 2).

Пусть в процессе градиентного спуска мы оказались в точке A, лежащей на поверхности f_1^*, ограничивающей область изменения переменных. Антиградиент минимизируемой функции $\left(-\frac{\partial З}{\partial X}\right)$, вычисленный в точке A, направлен за пределы допустимой области (вектор AA_1 на рис. 2). Допустимым направлением движения из точки A, соответствующим наибольшей скорости возрастания функции $З$, является направление вектора AA_2, совпадающего с проекцией антиградиента функции $З$ на плоскость L_1, перпендикулярную к вектору градиента функции ограничения f_1 в рассматриваемой точке A. Обозначим эту проекцию антиградиента Q. Следовательно, в данном случае движение производится в плоскости L_1, касательной в точке A к поверхности $f_1 = f_1^*$. Таким образом, при использовании метода проектирования градиента направление движения из точки, лежащей на границе допустимой области, определяется взаиморасположением вектора антиградиента функции $З(X)$ в этой точке и вектора градиента функции ограничений f_p в этой же точке.

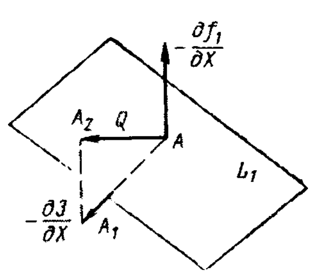

Рис.2. Графическая интерпретация метода проектирования градиента

Для более общего случая при наличии нескольких ограничивающих нелинейных неравенств движение из точки, расположенной на граничной поверхности f_ρ, препятствующей движению по направлению антиградиента $\left(-\frac{\partial З}{\partial X}\right)$, производится по одному из следующих вариантов:

1) по направлению проекции Q_ρ антиградиента минимизируемой функции $\left(-\frac{\partial З}{\partial X}\right)$ на соответствующую касательную плоскость L_ρ, если справедливо неравенство:

$$\left(Q\frac{\partial f_\rho}{\partial X}\right) > 0; \tag{13}$$

2) по направлению проекции вектора антиградиента функции $З(X)$ на пересечение касательных плоскостей, если неравенство (13) не выполняется.

При перемещении из точки A^i, находящейся на поверхности ограничений, на шаг конечной длины в одном из указанных направлений следующая точка поиска A^{i+1} может оказаться вне области допустимых значений переменных. Возвращение из точки A^{i+1} на поверхности ограничения производится движением по направлению $S=\sum_{\rho=1}^{N}\frac{\partial f_\rho}{\partial X}$, т. е. по направлению суммы градиентов функций ограничений, нарушенных в точке A^{i+1}. Заканчивая рассмотрение метода проектирования градиентов, следует отметить, что последовательности возможных движений при различных степенях ограничения весьма многочисленны и, очевидно, многократные проверки и проектирования при различных сочетаниях ограничений по F и Y приведут к весьма сложной программе счета, если придется иметь дело одновременно с несколькими ограничениями в виде неравенств. Метод проектирования градиента представляется весьма перспективным в тех случаях, когда уравнения ограничений линейны или почти линейны. К сожалению, в задачах оптимизации параметров теплоэнергетических установок это условие не выдерживается.

Для решения задач минимизации функции многих переменных при наличии ограничений в виде нелинейных неравенств можно также применить весьма простой в части алгоритма и программы метод штрафов. Суть его заключается в том, что в случае нарушения указанных ограничений к минимизируемой функции прибавляется некоторая положительная величина, подсчитанная как функция нарушенных ограничений. Тем самым такая система штрафов воздействует на направление изменения тех независимых переменных, которые привели к нарушению системы нелинейных неравенств. При выборе штрафной функции необходимо соблюдать следующие условия:

1) она должна быстро возрастать по мере нарушения ограничений в виде неравенств;

2) она должна быть вогнутой, так как иначе могут появиться посторонние решения (локальные минимумы за пределами допустимой области изменения параметров).

В остальном вид штрафной функции может быть произвольным, с оговоркой, что от ее вида зависит сходимость итерационных процессов. В настоящее время чаще всего применяется следующее задание штрафной функции:

$$З_{шт}=\frac{1}{2}\sum_{\rho=1}^{N}\xi_\rho k_\rho \Delta f_\rho^2, \tag{14}$$

где $\xi_\rho = 0$, если ограничения по f_ρ выполняются, и $\xi_\rho = 1$, если ограничения по f_ρ не выполняются; N – число нарушенных ограничений; k_ρ – положительная константа;

$$\Delta f_\rho = \begin{cases} f_\rho - f_\rho^{**} \text{ при } f_\rho > f_\rho^{**}; \\ f_\rho^{*} - f_\rho \text{ при } f_\rho < f_\rho^{*}. \end{cases}$$

Применительно к задачам оптимизации теплоэнергетических установок для определения штрафа можно также применять функцию, достаточно точно отражающую сущность происходящего при нарушении ограничения по f_ρ процесса. Примером такого подхода может быть точное определение эрозионного ущерба, наносимого запыленным потоком при выходе его скорости за максимально допустимое значение. Указанный подход к выбору выражения штрафной функции позволяет в процессе решения задачи оптимизации уточнить предельные значения функций ограничения (в приведенном примере – максимально допустимое значение скорости потока), принимаемые в ряде случаев, особенно для теплоэнергетических установок новых типов, сугубо ориентировочно.

В случае использования метода штрафных функций задачу минимизации функции $З(X)$ при наличии ограничений в форме нелинейных неравенств можно заменить задачей нахождения минимума более сложной функции, но лишенной ограничений в виде неравенств. Это новая минимизируемая функция имеет вид:

$$З' = З(X) + \frac{1}{2}\sum_{\rho=1}^{N} \xi_\rho k_\rho \Delta f_\rho^2. \qquad (15)$$

Совершенно очевидно, что «штрафные» слагаемые во втором члене выражения (15), будучи неотрицательными, при нарушении границы какой-либо функции f_ρ приводят к резкому увеличению функции $З'$ и тем самым препятствуют движению в недопустимую область.

К настоящему времени накоплен положительный опыт применения метода штрафных функций для решения ряда практических задач оптимизации. Вместе с тем в сложных задачах при большом числе нелинейных ограничений в виде неравенств, когда точка оптимума может лежать на границах нескольких из этих ограничений, применение способа штрафных функций дало недостаточно хорошие результаты. Дело в том, что неоднозначное изменение минимизируемой функции вследствие периодического появления или исчезновения отдельных функций штрафа приводит к систематическому, очень резкому изменению направлений антиградиента; при этом истинное направление спуска теряется, скорость спуска замедляется, а время решения на ЭВМ интенсивно растет. Иногда методом штрафов вообще не удается преодолеть «зацикливания» и получить решение задачи [2].

Изложенные методы учета нелинейных границ в процессе оптимизации основываются на итерационном возвращении в допустимую область после установления факта выхода за ее границы. Они дают удовлетворительные результаты при минимизации целевой функции в выпуклой области допустимых значений переменных. В случае выпукло–вогнутых ограничений, как это имеет место в рассматриваемом классе задач, использование такого приема нарушило бы сходимость процесса оптимизации. В самом деле, точки пересечения границ при итерационном возвращении к ним из недопустимой области согласно одному из изложенных методов могут не совпасть с точками пересечения границ при их нарушении по направлению скорейшего спуска. В результате этого значение минимизируемой функции после возврата $З^{i+1}$ может оказаться выше, чем значение до выхода из допустимой области $З^{i}$, что противоречит основной цели решаемой задачи. Поэтому при оптимизации параметров теплоэнергетических установок более целесообразно применение алгоритмов спуска точно до границ допустимой области (без их нарушения) и дальнейшего движения непосредственно вдоль нелинейных ограничивающих функций F и ограничивающих параметров X. Применительно к первым двум методам ввода в допустимую область такая модернизация алгоритмов достаточно легко осуществима путем корректировки шага.

Эффективным методом решения многих нелинейных задач является метод динамического программирования, не предъявляющий никаких требований к виду целевой функции, кроме требования аддитивности, т. е. представления в виде суммы однотипных составляющих на всех этапах процесса оптимизации. Основная идея метода заключается в разделении сложной многофакторной задачи на ряд последовательных шагов. В простейшем виде задача динамического программирования выглядит следующим образом: найти экстремум функции

$$\sum_{i=1}^{n} f_i(x_i)$$

при ограничениях

$$\sum_{i=1}^{n} x_i \le A.$$

Решение задачи производится на основе рекуррентных соотношений Беллмана:

$$h(c) = \min\left[h_{k-1}(c - x_k) + f_k(x_k)\right].$$

Вычислительная схема метода сводится к получению последовательности функций $h_k(x)$ *для* $k = \overline{1,n}$. На последнем шаге отыскивается оптимальное значение целевой функции, а значения переменных, удовлетворяющих оптимальному решению, находятся "обратным ходом".

Недостатками метода динамического программирования можно считать большой объем вычислений и необходимость хранения в памяти машины значительной по объему информации.

Получает распространение геометрическое (критеральное) программирование. В общем виде задача геометрического программирования формулируется следующим образом: минимизировать

$$f(x_1,\dots,x_m) = \sum_{i=1}^{n} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_m^{a_{im}} = \sum_{i=1}^{n} c_i \prod_{j=1}^{m} x_j^{a_{ij}}$$

при ограничениях

$$f_k(x_1,\dots,x_m) \le 1, \qquad k = \overline{1,r};$$
$$x_j \le 0; \quad j = \overline{1,m}.$$

Существенной особенностью геометрического программирования является то, что основное внимание акцентируется на слагаемых критерия оптимальности $c_i \prod_{j=1}^{m} x_j^{a_{ij}}$ и не предъявляется серьезных требований к виду оптимизируемой функции. Решение задач методом геометрического программирования базируется на понятии двойственности функции. Вследствие того, что метод геометрического программирования появился относительно недавно, он еще не получил широкого практического применения [1].

Таким образом, на основании вышеприведенного анализа можно сделать следующие выводы:

1. Все изложенные методы обладают определенными преимуществами и недостатками при оптимизации параметров теплоэнергетических установок. При использовании реальных

зависимостей между характеристиками теплоэнергетических установок, являющимися, как правило, нелинейными, следует использовать методы нелинейного программирования.

2. Методы пространственной сетки и Монте-Карло могут быть применены лишь для задач с небольшим числом оптимизируемых параметров теплоэнергетических установок. При решении задач с большим числом оптимизируемых параметров наиболее приемлемым является градиентный метод.

3. По сравнению с детерминированными методами метод случайного поиска обладает двумя важными преимуществами: во-первых, реализующие его алгоритмы менее чувствительны к наличию неглубоких локальных минимумов, во-вторых, в некоторых вариантах алгоритмы случайного поиска позволяют определять точку абсолютного минимума.

ЛИТЕРАТУРА

1. Яковлев Б.В. Повышение эффективности систем теплофикации и теплоснабжения. – Мн.: Адук.і вых., 2002. – 447 с.
2. Анциферов Е.Г., Ащепков Л.Т., Булатов В.П. Методы оптимизации и их приложения. – Новосибирск: Наука, 1990. – 159 с.
3. Попырин Л.С. Математическое моделирование и оптимизация теплоэнергетических установок. – М.:Энергия, 1978. – 416 с.
4. Тарасик В.П. Математическое моделирование технических систем. – Мн.: Дизайн ПРО, 2004. –639 с.