Стратегия оптимизационного исследования и методы решения задач статической и динамической оптимизации технологических объектов

Задачи статической оптимизации технологических объектов традиционно формулируются в форме задачи нелинейного программирования (НЛП) с ограничениями типа равенств и неравенств. В работах [45-50] установлено, что в случае многих переменных квадратичная аппроксимация (например используемая в методе Ньютона) обычно дает хорошие оценки точек безусловного минимума. Более того, группа квазиньютоновских методов позволяет пользоваться преимуществами квадратичной аппроксимации, не строя в явном виде полную аппроксимирующую функцию второго порядка на каждой итерации. Квазиньютоновские методы способны ускорить вычислительный процесс при использовании их в рамках процедур определений направлений поиска для методов приведенного градиента и проекций градиента.

В методе последовательного квадратичного программирования решение общей задачи НЛП ищется путем замены каждой нелинейной функции локальной квадратичной аппроксимацией в точке приближенного решения x^0 и решения получаемой последовательности аппроксимирующих подзадач. При этом установлено [46], что для задач квадратичного программирования существуют специальные методы, дающие решение за конечное число итераций без одномерного поиска при использовании вместо него итерации симплексного типа.

В 1980 г. К. Шитковский опубликовал в работе [50] результаты обширного исследования программ НЛП. В экспериментах использовались более 20 программ и 180 тестовых задач, генерируемых случайным образом; при этом структура задач была заранее определена, и для каждой их них многократно задавались начальные приближения. Тесты были проведены для четырех программ методов штрафных функций, 11 программ методов множителей Лагранжа, трех программ метода обобщенного приведенного градиента (ОПГ) и четырех программ метода последовательного квадратичного программирования (ПКП).

Программы оценивались по следующим критериям: 1) робастность; 2) надежность; 3) глобальная сходимость; 4) способность решать вырожденные и плохо обусловленные задачи; 5) чувствительность к малому изменению условий задачи; 6) простота обращения с программой.

На основе многочисленных тестов К. Шитковский пришел к весьма интересным выводам относительно классов алгоритмов и дал рекомендации по разработке программного обеспечения. В соответствии с его исследованиями классы алгоритмов можно проранжировать следующим образом: 1) методы ПКП; 2)Методы ОПГ; 3)методы множителей; 4) методы штрафных функций, не вошедшие в первые три класса.

Теперь рассмотрим некоторые принципы проведения оптимизационного исследования. Известно, что задача, к которой можно применить оптимизационные методы, должны включать критерий эффективности, независимые переменные, ограничения в виде равенств и неравенств, которые и образуют модель рассматриваемой системы.

Описанные и построенные модели реального объекта – важнейший этап оптимизационного исследования, так как он определяет практическую ценность получаемого решения и возможность его реализации.

Процесс оптимизации с использованием модели можно рассматривать как метод отыскания оптимального решения для реального объекта без непосредственного экспериментирования с самим объектом. “Прямой” путь, ведущий к оптимальному решению, заменяется “обходным”, включающим построение и оптимизацию модели, а также преобразование полученных результатов в практически реализуемую форму. Очевидно, что такой подход к оптимизации объекта обязательно требует использования некоторого упрощенного представления реального объекта. При формировании такого приближенного представления или модели следует учитывать только важнейшие характеристики объекта, которые должны быть отражены в модели, а менее существенные особенности в модель можно не включать. Необходимо также сформулировать логически обоснованные допущения, выбрать форму представления модели, уровень ее детализации и метод реализации на ЭВМ. Указанные соображения относятся к этапу построения модели и являются в той или иной мере произвольными. Модели можно упорядочить по степени адекватности описания поведения реального объекта в представляющей интерес области эксплуатации. Таким образом, качество модели нельзя оценивать ни по структуре, ни по форме. Единственным критерием такой оценки может служить лишь достоверность полученных на модели примеров поведения реального объекта.

В то же время адекватность модели часто невозможно строго оценить и поэтому выбор той или иной модели в значительной степени субъективен. Так, например, одна модель может оказаться более точной, чем другая, в определенном диапазоне изменения переменных, но менее точной в другом диапазоне.

Следует отметить, что соответствие модели реальному объекту носит в лучшем случае правдоподобный характер. Поскольку модель по своей сути не более чем упрощение реальных соотношений, то не существует абсолютных примеров, с помощью которых можно было бы ранжировать модели. Всегда есть ситуации, требующие субъективной оценки и предвидения того, как поведет себя реальный объект. Как следствие очень важно, чтобы создатель модели детально знал моделируемую систему, понимал технические принципы, лежащие в основе модели, а в случае оптимизации проекта сам руководил вычислениями, необходимыми для получения практически реализуемого проекта.

Работа по созданию модели является самым дорогим этапом оптимизационного исследования, так как она требует привлечения компетентных специалистов, хорошо знающих предметную область и изучаемый объект. Поскольку стоимость создания моделей резко возрастает по мере их детализации, необходимо тщательно продумывать уровень детализации, чтобы он соответствовал целям исследования и отвечал качеству доступной информации об объекте.

В оптимизационных исследованиях обычно используются модели трех основных типов: 1) аналитические; 2) модели поверхности отклика (регрессионные); 3) имитационные.

Вычислительные трудности, связанные с решением задачи, обычно вызываются четырьмя основными причинами: плохим масштабированием, несоответствием программ для вычисления значений функции и программ для вычисления производных, недифференцируемостью входящих в модель функций, неправильным заданием области определения значений аргументов функций. Только при тщательном анализе модели можно выявить эти ситуации и исключить их путем простой модификации модели.

В результате масштабирования осуществляется переход к относительным значениям величин, используемых в модели. В идеальном случае все переменные модели масштабируются таким образом, чтобы их значения находились в интервале $0.1-10$. Таким же образом по оценкам ограничений в приближенном решении исследуется чувствительность ограничений к изменениям значений переменных. Для этого вычисляется матрица, составляемая из градиентов ограничений. Наилучший случай, когда все ограничения имеют почти одинаковую чувствительность к изменениям значений переменных и значения градиентов ограничений находятся внутри одного и того же интервала значений. Благодаря этому невязки ограничений получают одинаковые веса и матричные операции с якобианом ограничений не приводят к потере точности вычислений.

Для надежной оптимизации объектов, целевые функции которых могут иметь несколько локальных минимумов, следует воспользоваться несколькими методами решения задачи, чтобы найти глобальный минимум. Отыскать глобальный минимум желательно не только в связи с тем, что это лучшее возможное решение задачи, но также и потому, что локальный минимум может провести к неправильным оценкам результатов расчетов по определению влияния переменных модели. Методы поиска глобального оптимума являются в настоящее время предметом интенсивных исследований. Известные методы поиска делятся на детерминированные и стохастические, которые в свою очередь могут быть эвристическими и строго обоснованными. Простейший и наиболее широко используемый метод состоит в проведении ряда оптимизационных расчетов при различных начальных условиях. В этом методе начальные точки выбираются из определенной решетки или же генерируются случайным образом. В первом случае допустимая область разбивается на непересекающиеся области и оптимизация выполняется каждой такой области по отдельности. Во втором случае начальные точки выбираются случайным образом, считая, что они распределены равномерно. В обоих случаях в качестве глобального оптимума из всех найденных локальных минимумов принимается локальный минимум с минимальным значением целевой функции. Оба этих метода эвристические. Теоретически, обратные методы глобальной оптимизации разработаны только для задач со специальной структурой .

Оптимизационные исследования не заканчиваются получением решения задачи. Напротив, самая важная часть исследования заключается в обосновании правильности решения и анализе его чувствительности. Наиболее важным является информация о состоянии объекта в окрестности решения, что позволяет глубже понять его основные свойства. Важнейшими результатами исследования являются ответы на такие вопросы: 1) какие ограничения активны в полученном решении? 2) что составляет основную часть затрат (стоимости)? 3) какова чувствительность решения к изменениям значений параметров?

Активные ограничения указывают на ограниченные возможности объекта или на то, что из-за проектных соображений объект усовершенствовать нельзя. По величине затрат (стоимости) находят тот блок объекта, параметры которого должны быть улучшены. Чувствительность решения к изменению значений параметров указывает на то, какие оценки параметров следует улучшить для того, чтобы безошибочно найти оптимально решение.

Рассмотренную выше стратегию оптимизационного исследования будем применять для решения задачи интегрированного проектирования технологических объектов и систем управления.

Далее остановимся на методах динамической оптимизации технологических объектов. Пусть функционирование управляемого технологического объекта (аппаратура, установки и т.п.) описывается на интервале $[t_1, t_2]$ дифференциальным уравнением

$$\dot{x}(t) = f(x,u,t), \ \ x \in E^n, n \in E^r. \tag{4.15}$$

Будем считать, что область допустимых управлений есть множество всех ограниченных кусочно-непрерывных функций $u(t)$ на $[t_0, t_1]$ таких, что $u \in U$ для любого $t \in [t_0, t_1]$, где $u \in E^r$ – заданное подмножество из r-мерного евклидова пространства E^r.

Введем скалярный критерий качества

$$I = V_3(x(t_1), t_1) + \int_{t_0}^{t_1} L(x,u,t)dt \tag{4.16}$$

где $L(x,u,t)$ – действительная функция на $E^n \times E^r \times [t_0, t_1]$ и $V_3(x(t_1), t_1)$ – действительная функция на $E^n \times [t_0, t_1]$.

Будем считать, что функции $f(x,u,t)$ и $L(x,u,t)$ непрерывны и дифференцируемы по совокупности переменных x,u,t. Пусть S – заданное множество из $E^n \times [t_0, t_1]$, назовем S множеством целей (множеством конечных состояний) и $V_3(x(t_1), t_1)$ – функцией конечных состояний.

Задачей оптимального управления для системы (4.15) при сделанных предположениях относительно начального состояния $x(t_0) \in E^n$, области $u \in E^r$ допустимых управлений $u(t) \in U$ и множества конечных состояний S является отыскание такого управления $u(t) \in U$, что функционал (4.16) достигает минимального значения.

Конкретизация выражений $f(x,u,t)$, $L(x,u,t)$, $V_3(x(t_1), t_1)$ и множества целей S порождает различные типы задач оптимального управления].

Классическое вариационное исчисление (в случае непрерывности $u(t)$) и принцип максимума Л.С. Понтрягина сводят задачу оптимального управления к решению двухточечной краевой задачи для системы нелинейных дифференциальных уравнений. Принцип максимума применим к задачам с управлением общего вида. В случае описания движения объекта линейными дифференциальными уравнениями общая теория задач оптимального управления, основанная на проблеме моментов, предложена и обоснована Н.Н. Красовским [51].

Характерным для задач оптимального управления является то, что точные аналитические решения удается получить лишь в редких случаях. К этим случаям относятся задачи с линейными объектами и квадратичными функционалами.

Сложность или невозможность получения аналитических результатов для задач в более общей постановке привели к развитию вычислительных и приближенных методов построения оптимального управления [52].

Решение сформулированной выше задачи оптимального управления получают обычно в форме так называемого программного управления, т.е. $u^* \equiv u^*(t)$, которое реализуется в разомкнутой системе управления. Применение таких систем управления процессами химической технологии не дает желаемого результата ввиду больших затрат машинного времени для расчета программы управления из-за изменчивости начальных условий и неточности реализации программы в процессе его функционирования. В связи с этим более перспективным направлением в автоматизации и оптимизации динамических режимов процессов химической технологии является синтез систем управления с обратной связью.

Методы аналитического конструирования оптимальных регуляторов (АКОР) позволяют синтезировать оптимальный закон управления (оператор обратной связи в виде $u^* = \psi(x)$ [53-56].

Остановимся здесь на модифицированном А.А. Красовским [54] методе аналитического конструирования, который заключается в видоизменении минимизируемого функционала, позволяющим численно получить решение для достаточно сложных нелинейных задач динамической оптимизации. Пусть управляемый процесс описывается дифференциальным уравнением типа

$$\dot{x} = f(x,t) + \varphi(x,t)\cdot u\,,$$

а минимизируемый функционал имеет вид

$$I = V_3[x(t_1)] + \int_{t_0}^{t_1} Q_3[x(t),t]dt + \int_{t_0}^{t_1} \{U_3[u(t),t] + U_3^*[u^*(t),t]\}dt, \tag{4.17}$$

где U_3, U_3^* – заданные функции аргументов такие, что $\left\{U_3(u,t) + U_3^*(u^*,t) - \left[\frac{\partial}{\partial u}U_3(u^*,t)\right]\cdot u\right\}$ – положительно определенная функция относительно u, обращающаяся в нуль при $u = u^*$. Заметим, что функция u^* в (4.17) – пока неизвестное оптимальное управление.

В работе [54] показано, что оптимальное управление $u = u^*$ в данном случае определяется соотношением

$$\frac{\partial U_3(v,t)}{\partial V} = -\frac{\partial V}{\partial x}\varphi(x,t)\,,$$

где $V = V(x,t)$ есть решение уравнения Ляпунова для неуправляемого ($u \equiv 0$) объекта

$$\frac{\partial V}{\partial t} + -\frac{\partial V}{\partial x} f(x,t) = -Q_3(x,t)$$

при граничном условии

$$V_{t=t_1} = V_3(x)\,.$$

Для случая функционала (4.17) с квадратичной функцией

$$V_3 = \frac{1}{2}\sum_{j=1}^{r}\frac{u_j^2 + u_j^{*2}}{k_j^2}dt$$

оптимальным управлением являются функции

$$u_j = u_j^* = -k_j^2\sum_{i=1}^{n}\varphi_{ij}(x,t)\frac{\partial V}{\partial x_i}, \quad j = \overline{1,r}\,. \tag{4.18}$$

Таким образом, оптимальное управление при функционале “обобщенной работы” А.А. Красовского (4.17) имеет такой же внешний вид, как и при классическом функционале. Однако функция $V = V(x,t)$ здесь есть решение линейного уравнения с частными производными

$$\frac{\partial V}{\partial t} - \sum_{i=1}^{n}\varphi_j\frac{\partial V}{\partial x_i} = -Q_3\,. \tag{4.19}$$

при граничном условии

$$V_{t=t_1} = V_3, \tag{4.20}$$

в то время как при классическом функционале $V = V(x, t)$ есть решение нелинейного уравнения Беллмана. Это принципиальное отличие, сохраняющееся для всех задач оптимального управления по функционалу обобщенной работы, обусловливает широкие возможности для синтеза систем оптимального управления периодическими процессами и пусковыми режимами непрерывных процессов химических производств.