

Решение оптимизационных задач методом Ньютона с применением быстрого автоматического дифференцирования

Е.С.Засухина

Вычислительный центр РАН, zasuhina@ccas.ru

К числу наиболее распространенных и плодотворных методов, к которым прибегают при решении оптимизационных задач, относятся градиентные методы. Еще более привлекательными в смысле нахождения решения с высокой точностью, являются методы, использующие вторые производные. При этом точность и время вычисления производных существенно влияют на эффективность алгоритмов оптимизации в целом.

В ВЦ РАН под руководством Ю.Г. Евтушенко была разработана методология быстрого автоматического дифференцирования (БАД-методология), позволяющая с единых позиций определять градиенты не только функций, заданных в явном виде, но и функций, не имеющих явного аналитического представления, например, функций, значения которых определяется в результате выполнения некоторого многошагового процесса. Примером могут служить функции, возникающие в результате дискретной аппроксимации задач оптимального управления процессами, описываемыми обыкновенными дифференциальными уравнениями или уравнениями с частными производными. Этот подход был успешно применен при решении ряда задач оптимального управления.

Однако, с ужесточением требований на точность нахождения решения, а также, как хорошо известно, в случае "овражного" вида минимизируемой функции возникает настоятельная необходимость применения в вычислительном процессе метода Ньютона.

Предлагаемая работа распространяет БАД-методологию на случай вычисления вторых производных. В работе выводятся точные формулы для вычисления вторых производных сложных функций, на переменные которых наложены функциональные связи. Эти формулы содержат сопряженные переменные, которые определяются в результате решения сопряженного матричного уравнения. Установлено, что основная матрица этого уравнения и основная матрица сопряженной системы, которую необходимо решить при определении градиента функции, являются транспонированными друг другу. Количество скалярных уравнений, которые должны быть решены для определения вторых производных функции, есть линейная функция размерности независимой переменной.

Полученные формулы адаптируются для случая многошаговых процессов, получаемых в результате дискретизации задачи оптимального управления процессом, описываемым системой обыкновенных дифференциальных уравнений с помощью метода Рунге-Кутты произвольного порядка. В работе исследуется структура возникающих при этом сопряженных систем. Показано, что эта структура оказывается достаточно простой. Указывается метод решения сопряженных систем.

Формула для вычисления вторых производных записывается в виде, позволяющем выделить множество пар слагаемых, транспонированных друг другу. На основе использования особенностей структуры этих слагаемых, их разреженности, симметричности матрицы вторых производных, строится алгоритм, сокращающий объем вычислений и требуемую память.

С целью оценить эффективность предлагаемого подхода были проведены две группы расчетов. Первую из них составляют расчеты, в которых оценивалось отношение времени, затрачиваемого на вычисление вторых производных сложной функции, ко времени вычисления самой функции. Для этого рассматривались три задачи оптимального управления процессами, описываемыми системами обыкновенных дифференциальных уравнений. Эти задачи дискретизировались по схеме Эйлера, для полученных задач вычислялись целевые функции и гессианы. Проводились замеры времени, затрачиваемых на вычисление целевой функции и гессиана. Расчеты проводились для сеток с различным количеством узлов, чтобы оценить, как изменяется отношение времени вычисления гессиана t_{hess} ко времени вычисления t_f при изменении размерности r переменной управления и как это отношение изменяется при переходе от одной задачи к другой. Расчеты показали, что

значение $\frac{t_{hess}}{(t_f \ r)}$ незначительно изменяется в зависимости от r , но существенно зависит от конкретной

задачи. Таким образом, полученные результаты позволяют сделать заключение о том, что время, затрачиваемое на вычисление вторых производных целевой функции, по отношению ко времени, затрачиваемому на вычисление самой функции, с увеличением r растет линейным образом.

Вторую группу составили расчеты с помощью метода Ньютона конкретной задачи оптимального управления, дискретизированной методом Рунге-Кутты 1-го, 2-го, 3-го, 4-го порядков. С целью оценить эффективность метода Ньютона, использующего предлагаемый алгоритм для вычисления вторых производных, расчеты проводились и градиентным методом. Расчеты производились с хорошим начальным приближением и с произвольным начальным приближением. В случае произвольного начального приближения использовался модифицированный метод Ньютона, в котором длина шага определялась как минимум функции, полученной в результате интерполяции целевой функции вдоль выбранного направления с помощью кубических сплайнов. Кроме того, осуществлялся контроль за тем, чтобы каждая последующая итерация

приводила бы к убыванию целевой функции. Расчеты проводились с использованием различных количеств точек, образующих интерполирующий сплайн.

Анализ результатов расчетов показал, что независимо от порядка аппроксимации и от количества точек, определяющих интерполирующий сплайн, метод Ньютона демонстрирует преимущество перед градиентным методом по всем показателям, включая и время.