Численные методы оптимизации

Занятие 15
декабрь 2011

кафедра систем телекоммуникаций
Ловецкий К.П.

Стандартная постановка задачи конечномерной оптимизации (P)

Основная задача: Минимизировать целевую функцию

$F(x): R^n \to R$ на допустимом множестве $C \subset R^n$.

$$(P) \qquad F(x) \to \min \qquad x = (x_1, \ldots, x_n) \in X \subset R^n$$

$$c_i(x) \geq 0, \quad i = 1, \ldots, m$$

Введем обозначения:

$$g_i(x) = \frac{\partial F(x)}{\partial x_i}, \quad i = 1, 2, \ldots, n$$

$$a_{ik}(x) = \frac{\partial c_i(x)}{\partial x_k}, \quad i, k = 1, 2, \ldots, n$$

10.03.2012 2

Условия первого порядка

- **Теорема 11.1.** Пусть точка x - решение задачи условной минимизации функции $F(x)$. Тогда не существует вектора p, для которого выполняется неравенство

$$g(x)^T p < 0$$

и при любых значениях α из отрезка $0 \le \alpha \le \bar{\alpha}, \bar{\alpha} > 0$, гарантирована допустимость точек вида $x + \alpha p$.

- **Лемма Фаркаша.** Пусть заданы векторы $a_i \quad (i = 1, ..., t)$ и вектор $g(x)$. Неравенства

$$g(x)^T p < 0$$

и

$$a_i^T p \ge 0, i = 1, ..., t,$$

несовместны тогда и только тогда, когда $g(x)$ принадлежит выпуклому конусу, натянутому на векторы $a_i \quad (i = 1, ..., t)$.

Условия первого порядка

- **Теорема 11.2.** Пусть x^* - точка, доставляющая минимум функции $F(x)$ при линейных ограничениях $a_i^T x \ge b_i, i = 1,...,m,$ причем первые из них обращаются в x^* в равенства. Тогда градиент $g(x^*)$ представим в виде

$$g(x^*) = \sum_{i=1}^{t} \lambda_i a_i, \quad \lambda_i \ge 0.$$

- **Теорема 11.3.** Пусть x^* - точка, доставляющая минимум функции $F(x)$ при нелинейных ограничениях $c_i(x) \ge 0, \quad i = 1,...,m$ причем первые t из них обращаются в x^* в равенства и градиенты $a_i(x^*),(i = 1,...,t)$ линейно независимы. Тогда градиент $g(x^*)$ представим в виде

$$g(x^*) = \sum_{i=1}^{t} \lambda_i a_i(x^*), \quad \lambda_i \ge 0.$$

Функции Лагранжа

- Итак, как показано ранее, в точке минимума функции $F(x)$ при ограничениях $c_i(x) \geq 0, \quad i = 1,...,m$, обычно выполняется соотношение

$$g(x^*) = \sum_{i=1}^{t} \lambda_i a_i(x^*), \quad \lambda_i \geq 0.$$

где $g(x)$ и $a_i(x)$ - градиенты функций $F(x)$ и $c_i(x)$. В предположении линейной независимости векторов $a_i(x^*), (i = 1,...,t)$ это разложение существует, причем множители λ_i определяются однозначно. Поэтому решение исходной задачи при соответствующем выборе вектора λ будет стационарной по x точкой функции

$$L(x, \lambda) = F(x) - \sum_{i=1}^{t} \lambda_i c_i(x).$$

Она называется функцией Лагранжа, а параметры λ_i - множителями Лагранжа.

Точная штрафная функция Флетчера

В методах обычных штрафных функций решение задачи определяется как последовательность решений подзадач безусловной минимизации. В связи с этим возникает желание построить *точную штрафную функцию* $\varphi(x)$ с локальным безусловным минимумом в точке x^*, что позволит решать задачу минимизации лишь один раз. При этом необходимо, чтобы функция была гладкой в окрестности точки x^*, а матрица $\nabla^2\varphi(x^*)$ была бы положительно определенной. В этом случае значение x^* можно найти посредством однократной безусловной минимизации функции $\varphi(x)$ при помощи одного из стандартных методов поиска минимума без ограничений.

Точная штрафная функция Флетчера

Если представить функцию $\varphi(x)$ как функцию Лагранжа, в которой вектор λ, состоящий из множителей Лагранжа, зависит от x, то расчет $\lambda(x)$ потребует конечного числа элементарных операций над значениями функций задачи и их производных в точке x.

Для того чтобы искомое решение x^* было стационарной точкой соответствующей функции $\varphi(x)$, т.е. $\nabla\varphi(x^*)=0$, достаточно, чтобы $\lambda(x)\to\lambda^*$ при $x\to x^*$. Другими словами, если

$$\varphi(x)=F(x)-\lambda(x)^T c(x)$$

то

$$\nabla\varphi(x)=g(x)-A(x)\lambda(x)-\left[\nabla\lambda(x)\right]c(x)$$

Точная штрафная функция Флетчера

$$\varphi(x) = F(x) - \lambda(x)^T c(x)$$

$$\nabla\varphi(x) = g(x) - A(x)\lambda(x) - \left[\nabla\lambda(x)\right]c(x)$$

где $A(x)$ - матрица, столбцами которой являются производные от вектор-функции ограничений $c_i(x)$ и, следовательно, равенство $\nabla\varphi(x^*) = 0$ получается при $x \to x^*$ в силу непрерывности g, A, λ, c и равенств $g^* = A\lambda^*, c^* = 0$. В качестве функции $\lambda(x)$, сходящейся к λ^* при $x \to x^*$, можно взять

$$\lambda(x) = A^+ g\Big|_x, \qquad A^+ = \left(A^T A\right)^{-1} A^T$$

где

то есть $\lambda(x)$ вычисляется как вектор, минимизирующий сумму квадратов невязок уравнений переопределенной системы $A\lambda = g$

Точная штрафная функция Флетчера

Тогда $\varphi(x) = F(x) - c(x)^T A^{+}(x) g(x)$

Посмотрим, будет ли x^* доставлять локальный минимум этой функции. Матрица ее вторых производных в точке x^* выглядит так

$$\nabla^2 \varphi\left(x^*\right) = \hat{P} W^* \hat{P} - P W^* P,$$

где $P = A^* A^{*+}$ - матрица проектирования, $\hat{P} = I - P$,

и $W^* = \nabla^2 F^* - \sum_i \lambda_i^* \nabla^2 c_i^*$

По поводу положительной определенности матрицы вторых производных сказать ничего нельзя $?\perp PW^* P,$ поэтому модифицируем исходную функцию:

$$\varphi(x) = F(x) - c(x)^T A^{+}(x) g(x) + q c(x)^T A^{+}(x) A^{+T}(x) c(x)$$

$$\nabla^2 \varphi\left(x^*\right) = \hat{P} W^* \hat{P} + P(2qI - W^*)P, \quad q > 1/2\left|W^*\right|$$

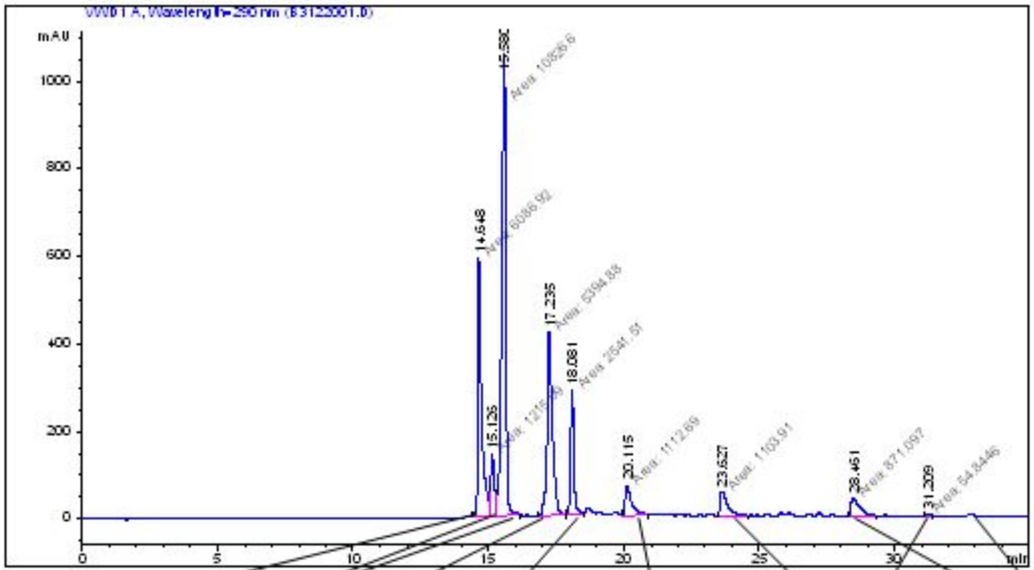
Пример.

- **Аппроксимация кривой гауссианами.**

$$\varphi(x)=\sum_{k=1}^{N}\left(a_0+\sum_{i=1}^{n}a_i\cdot e^{\left(\frac{-\left(x_k^{изм}-x_i\right)^2}{b_i}\right)}-F_k^{изм}\right)^2$$

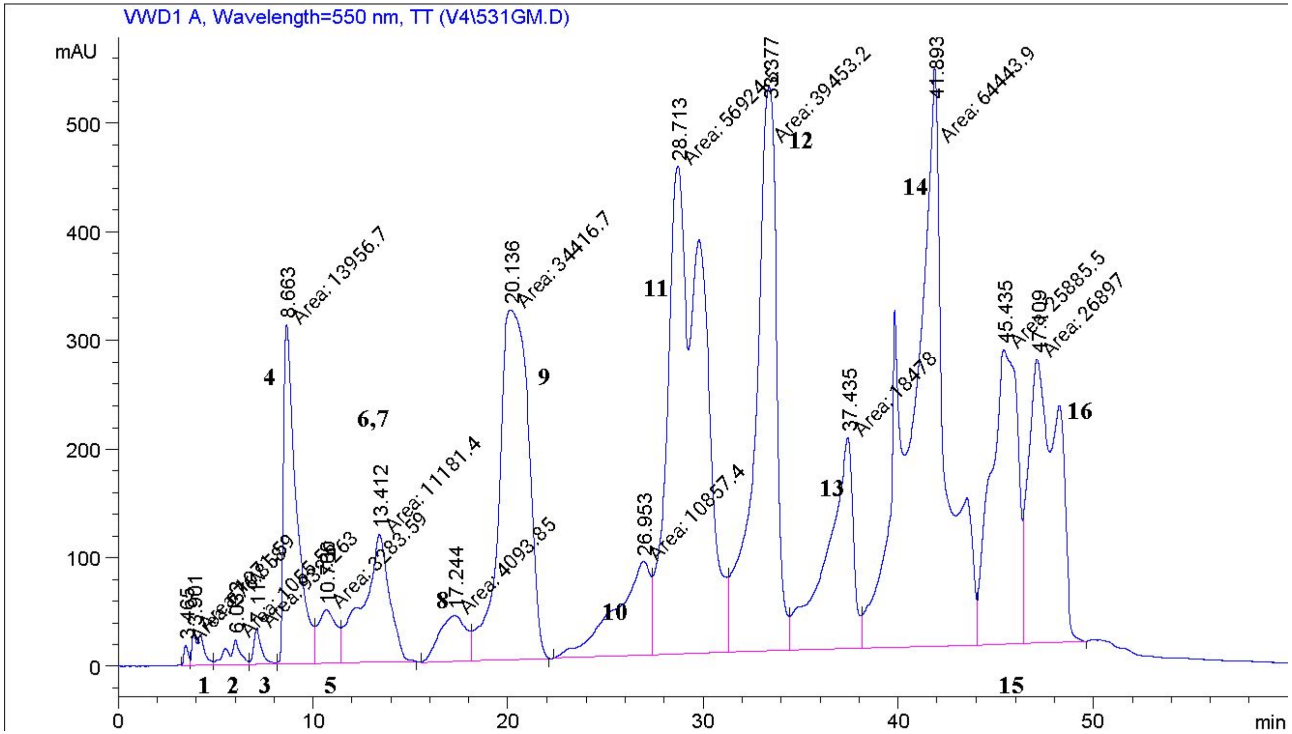
где N – количество измеренных точек спектра, n - количество функций Гаусса, a_i и b_i - параметры, задающие амплитуду и ширину гауссианов, $a_i>0, b_i>0$ x_i , $x_k^{изм}$ - координата середины гауссиана на оси абсцисс, $F_k^{изм}$ - значение спектра в измеренной точке

Основные компоненты синего красителя



1	2 -3	4	5	6	7 - 9	8 - 10
Тетра сульфо	Трисульфо	Дисульфо	Трисульфо хлор	Дисульфо хлор	Моносульфо	Моносульфо хлор
Компоненты высокой полярности					Компоненты низкой полярности	

Фиолетовый краситель



Результаты хроматографического анализа фиолетового красителя

Окно программы **Spectrum Analysis**