

Ю. В. Заика (Петрозаводск, ИПМИ КарНЦ РАН). **Алгоритм минимизации энергии Гиббса: расчет фазового равновесия.**

Задача определения равновесного состава газовой смеси является одной из традиционных в химической термодинамике. обстоятельный анализ и библиография содержится, например, в [1]. Что касается эффективных численных методов, то, по видимому, отчет следует вести от работы [2] (с участием автора симплекс-метода). Чтобы сосредоточиться на основной идее и использовать минимальные математические средства, остановимся на классической постановке задачи:

$$g(n) \equiv g(n_1, \dots, n_k) = \sum_{i=1}^k n_i [c_i + \ln(n_i/\bar{n})] \rightarrow \min, \quad \bar{n} \equiv \sum_{i=1}^k n_i = \|n\| \quad (n_i \geq 0, n \neq 0).$$

Здесь c_i — постоянные, n_i — количество молей i -й компоненты смеси. В дальнейшем считаем $c_i < 0$, но это не принципиально. Материальный баланс выражается ограничением $A'n = b > 0$. Элементы матрицы $A_{k \times m}$ — неотрицательные целые числа, нулевые строки отсутствуют, $\text{rank } A = m < k$. В квадратных скобках — компоненты градиента g , $(\text{grad } g)_i \rightarrow -\infty$ при $n_i \rightarrow +0$ ($0 < \bar{n}_{\min} \leq \bar{n} \leq \bar{n}_{\max} < +\infty$). Цель работы — учесть особенность ($\ln \rightarrow -\infty$), которая неизбежно сказывается на работоспособности методов оптимизации, основанных на использовании градиента и гессиана.

Используя мольные доли $x_i = n_i/\bar{n}$, запишем задачу в форме

$$g(n) = \bar{n}f(x) \equiv \bar{n}[c'x + \varphi(x)] \rightarrow \min, \quad A'n = b, \quad x \in S = \{x | x_1 + \dots + x_k = 1, x_i \geq 0\},$$

$\varphi(x) \equiv \sum x_i \ln x_i$. По непрерывности доопределяем $x_i \ln x_i = 0$ при $x_i = 0$. Экстремумы $\varphi(x)$ на S : $\max \varphi = 0$, $\min \varphi = -\ln k$. Максимум достигается, когда смесь вырождается в компоненту. Минимум единственный, $x_i = 1/k$. Целевая функция имеет двустороннюю оценку: $L_0(n) \leq g(n) \leq L^0(n)$, $L_0 = c'n - \bar{n} \ln k$, $L^0 = c'n$. Минимум $g_* = \min g$, $n \in D = \{n | n \geq 0, A'n = b\}$, оценивается решением двух задач ЛП.

Начальное приближение. Фиксируем единичный направляющий вектор $n^0 \in S$ ($\bar{n}^0 = 1$). На луче $n(t) = tn^0$ ($t \geq 0$ интерпретируем как время движения) имеем $g(n(t)) = tf(n^0)$. Естественно рассмотреть направление n^0 , вдоль которого g убывает наискорейшим образом: $f(n^0) \rightarrow \min$, $n_i^0 = x_i^0 = [\exp\{c_i - c_1\} + \dots + \exp\{c_i - c_k\}]^{-1}$. Теперь примем во внимание ограничение $A'n = b$. Пусть n^1, n^2 — решения задач ЛП $\bar{n} \rightarrow \min$, $\bar{n} \rightarrow \max$, $n \in D$. Помимо $n^{1,2}$ рассмотрим точки n^{3-6} экстремумов оценочных L_0, L^0 . Минимум из минимумов g на невырожденных отрезках $[n^i, n^j]$ определяет направление $\tilde{n}^0 \in S$. Целесообразно добавить равномерное направление $r^0 = (1/k, \dots, 1/k)'$ (стабилизатор) и допустимыми направлениями спуска h считать выпуклые комбинации n^0, \tilde{n}^0, r^0 . Фиксируем h и спроектируем движущуюся точку th , $t \geq 0$, на $\Lambda = \{z | A'z = b\}$: $p(t) = tH + B$, $H = h - A(A'A)^{-1}A'h$, $B = A(A'A)^{-1}b$. Векторное ограничение $p(t) \geq 0$ дает отрезок $[t_1, t_2]$, на котором $p(t) \in D$. В качестве первого приближения n_*^1 берем точку минимума выпуклой функции $g(p(t))$, $t \in [t_1, t_2]$.

Итерационный процесс. Ориентируясь на текущую (s -я итерация) локальную аппроксимацию $g(n) = c'n + \bar{n}\varphi(n) \approx c'n + \bar{n}\varphi(n_*^s)$, рассмотрим задачу ЛП $L_s^-(n) = c'n + \bar{n}\varphi(n_*^s) = \sum [c_i + \varphi(n_*^s)]n_i \rightarrow \min$, $n \in D$. Считаем $\varphi(0) = 0$, $\varphi(n) = \varphi(n/\bar{n}) = \varphi(x)$. По сравнению с «обычным» вариантом $L_s^+(n) = \langle \text{grad } g(n_*^s), n \rangle = \sum [c_i + \ln(x_*^s)]n_i \rightarrow \min$ диапазон изменения коэффициентов аппроксимирующей линейной формы сужается с несобственного множества $[-\infty, c_i]$ до равномерной поправки $\varphi(x_*^s) \in [-\ln k, 0]$. Попытаемся расширить возможности аппроксимации, объединив невырожденность L_s^- с экстремальными свойствами L_s^+ . Ограничим множество возможных значений коэффициентов линейной формы, поставив барьер неограниченному росту нормы градиента. Для этого нужен масштаб скорости. Естественно обратить внимание на оценку $L_0(n) \leq g(n) \leq L^0(n)$. Разумно фиксировать максимальную

по абсолютной величине скорость убывания $V = c_k - \ln k$ ($c_k = \min c_i < 0$) и не позволять недоминирующим компонентам двигаться существенно быстрее (локально, в линейном приближении). При классической линеаризации (форма L_s^+) в слагаемых целевой функции $[c_i + \ln x_i]n_i$ текущее приближение используется для фиксации нелинейности, зависящей явно лишь от мольной доли (замена $x_i \rightarrow x_{*i}^s$). Будем придерживаться этой схемы с той лишь разницей, что вследствие угрозы $\ln x_i \rightarrow -\infty$ будем сначала выделять функции $x_i \ln x_i$ (разрешимую особенность), оставляя n_i свободными переменными после подстановки $x_{*i}^s \rightarrow x_i$. Формализуем приведенные соображения.

Определим вектор \tilde{c} коэффициентов линейной аппроксимирующей формы $L_s(n) = \tilde{c}'n$ алгоритмически. Предварительно обнуляем массив: $\tilde{c} = 0$. Если $x_*^s = n_*^s/\bar{n}_*^s = e^j = (0, \dots, 1, \dots, 0)'$ (вырожденное приближение смеси одной компонентой), то полагаем $\tilde{c} = c$. При этом $L_s(n) = c'n = L_s^-(n) = L^0(n)$ — верхняя оценка целевой функции $g(n)$. Классическая форма $L_s^+(n)$ не имеет смысла ($\ln = -\infty$). Далее считаем $x_{*i}^s \neq 1$.

Шаг 1. Рассмотрим первое слагаемое $g(n)$, выделив явно разрешимую особенность: $n_1[c_1 + \ln x_1] = n_1c_1 + \bar{n}x_1 \ln x_1$, $x_1 = n_1/\bar{n}$, $\bar{n} = \sum n_i$. Если $x_{*1}^s = 0$ ($< \varepsilon \ll 1$), то, имея в виду подстановку $x_{*1}^s \rightarrow x_1$, полагаем $\tilde{c}_1 = c_1$. Пусть $x_{*1}^s > 0$ ($> \varepsilon$). Если выполнено неравенство $\xi_1 \equiv V[c_1 + \ln x_{*1}^s]^{-1} \geq 1 \Leftrightarrow (\text{grad } g(n_*^s))_1 \geq V$, то присваиваем $\tilde{c}_1 := c_1 + \ln x_{*1}^s$ (как и в L_s^+). Остается случай $\xi_1 \in (0, 1)$. В тождество $n_1[c_1 + \ln x_1] = \zeta n_1[c_1 + \ln x_1] + (1 - \zeta)[n_1c_1 + \bar{n}x_1 \ln x_1]$ подставим значения $\zeta = \xi_1^2$, $x_1 = x_{*1}^s$. Параметр ζ подбирается из соображений согласования при предельных переходах: $x_{*1}^s \rightarrow +0 \Rightarrow \tilde{c}_1 \rightarrow c_1$, $\xi_1 \rightarrow 1 - 0 \Rightarrow \tilde{c}_1 \rightarrow V$ (для этих целей допустимы и $\zeta_1 = \xi_1^{1+\nu_1}$, $\nu_1 > 0$). Получаем выражение $\xi_1 V n_1 + (1 - \xi_1^2)[n_1c_1 + \bar{n}x_{*1}^s \ln x_{*1}^s]$. При переменной n_1 имеем множитель $\tilde{c}_1 = \xi_1 V + (1 - \xi_1^2)[c_1 + x_{*1}^s \ln x_{*1}^s]$. Дополнительно из-за множителя \bar{n} следует для $j = 2, \dots, k$ переопределить $\tilde{c}_j := \tilde{c}_j + (1 - \zeta)x_{*1}^s \ln x_{*1}^s$ (до шага 1 $\tilde{c}_j = 0$). Итак, коэффициент \tilde{c}_1 определен однозначно.

Шаг 2. Переходим к следующему слагаемому $n_2[c_2 + \ln x_2] = n_2c_2 + \bar{n}x_2 \ln x_2$. Если $x_{*2}^s = 0$ ($< \varepsilon \ll 1$), то к определенному на шаге 1 значению \tilde{c}_2 добавляем величину c_2 : $\tilde{c}_2 := \tilde{c}_2 + c_2$. Пусть $x_{*2}^s > 0$ ($> \varepsilon$). Если $\xi_2 \equiv V[c_2 + \ln x_{*2}^s]^{-1} \geq 1$ (скорость убывания медленнее принятого порога), то добавляем к \tilde{c}_2 (на шаге 1) число $c_2 + \ln x_{*2}^s$. Остается случай $\xi_2 \in (0, 1)$. В тождество $n_2[c_2 + \ln x_2] = \zeta n_2[c_2 + \ln x_2] + (1 - \zeta)[n_2c_2 + \bar{n}x_2 \ln x_2]$ подставляем $\zeta = \xi_2^2$, $x_2 = x_{*2}^s$: $\xi_2 V n_2 + (1 - \xi_2^2)[n_2c_2 + \bar{n}x_{*2}^s \ln x_{*2}^s]$. К значению \tilde{c}_2 (определенному на шаге 1) добавляем $\Delta \tilde{c}_2 = \xi_2 V + (1 - \xi_2^2)[c_2 + x_{*2}^s \ln x_{*2}^s]$. Кроме того, к \tilde{c}_1 и \tilde{c}_j , $j = 3, \dots, k$, прибавляем число $(1 - \xi_2^2)x_{*2}^s \ln x_{*2}^s$. Аналогичные выкладки для варианта $\zeta_2 = \xi_2^{1+\nu_2}$, $\nu_2 > 0$.

Шаги 3, ..., k выполняются по аналогичной схеме преобразований слагаемых $n_i[c_i + \ln x_i] = n_i c_i + \bar{n} x_i \ln x_i$ в целевой функции g .

Заметим, что помимо $L_s = L^0$ ($\exists x_{*j}^s = 1$) реализуется и нижняя оценка g : $x_*^s = (1/k, \dots, 1/k)'$ $\Rightarrow L_s = L_s^- = L_0$. Когда все $\xi_i \geq 1$ (принятый ориентир $|V|$ скорости убывания не превышен), то $L_s = L_s^+ = \langle \text{grad } g(n_*^s), n \rangle$. Неограниченное убывание \tilde{c}_i исключено, поскольку особенность $\alpha \ln \alpha$ выделена явно и ограничена. Формирование \tilde{c} обобщенно можно считать процедурой регуляризации $\text{grad } g$. Далее переход от текущего приближения $n_*^s \approx n_*$ к следующему стандартен. Решение \tilde{n}_*^{s+1} задачи ЛП $L_s(n) = \tilde{c}'n \rightarrow \min$, $n \in D$, соединяем отрезком с вектором n_*^s . Приближение n_*^{s+1} определяется минимумом выпуклой функции $\psi(\lambda) = g(\lambda \tilde{n}_*^{s+1} + (1 - \lambda)n_*^s)$, $\lambda \in [0, 1]$.

Начальный спуск целесообразно осуществлять сразу по нескольким направлениям h . Алгоритм не использует квадратичной аппроксимации целевой функции. Его можно реализовать как блок генерации начального приближения в алгоритмах более высокого порядка (но обычно требующих хорошего начального приближения с положительными компонентами) с целью избежать вырождения вдали от минимума.

Работа выполнена при поддержке РФФИ, проект № 09-01-00439.

СПИСОК ЛИТЕРАТУРЫ

1. Уэйлес С. Фазовые равновесия в химической технологии. М.: Мир, 1989, Ч. 1,2.
2. White W. B., Johnson S. M., Dantzig G. B. Chemical Equilibrium in Complex Mixtures. — J. Chemical Physics, 1958, v. 28, № 5, p. 751–755.