Панферов С.В.

ГАРАНТИРОВАННАЯ ОЦЕНКА РАСПОЛОЖЕНИЯ ПЛОСКОСТИ КАСАТЕЛЬНОЙ К ОГРАНИЧЕНИЯМ В ОПТИМИЗАЦИОННОЙ ЗАДАЧЕ

Введение

Среди оптимизационных задач исследования операций одно из центральных мест занимают задачи с ограничениями. Для таких задач важна разработка численных методов. Большая группа методов – методы приведенного градиента [1] (Глава 2, стр.46). В этих методах градиент приводится относительно плоскости, касательной к ограничениям, активным в данной точке [2] (Глава 4, стр.132). Однако при практической реализации удобно использовать вместо касательных плоскостей координатные плоскости. Теоретическое обоснование подобной замены требует оценки величины угла между касательной и координатной плоскостями. Получению соответствующих гарантированных оценок посвящена данная работа.

Постановка задачи

Введем необходимые обозначения. Пусть $\mathbf{R}^k$ – k-мерное евклидово пространство $x=(x_1,x_2,\dots,x_k)$, $y=(y_1,y_2,\dots,y_k) \in \mathbf{R}^k$, $\langle x,y\rangle=\sum_{i=1}^{k}x_iy_i$ – скалярное произведение в $\mathbf{R}^k$, $\|x\|=\langle x,x\rangle^{1/2}$ – норма в $\mathbf{R}^k$. Пусть $e=(e_1, e_2, \dots, e_k)$ – ортонормированный базис (ОНБ) пространства $\mathbf{R}^k$, $L(h)=L(h_1, h_2,\dots,h_m)$ – линейная оболочка векторов $h_1, h_2,\dots,h_m \in \mathbf{R}^k$. Для произвольной m-мерной плоскости H в $\mathbf{R}^k$ через $L(H)$ будем обозначать параллельную H плоскость, проходящую через начало координат. Через $I=(i_1, i_2,\dots,i_m)$ обозначим набор индексов базисных векторов, задающих m-мерную координатную плоскость в $\mathbf{R}^k$, через $I_k^m=\{I=(i_1,i_2,\dots,i_m) \mid 1\le i_1<i_2<\dots<i_m\le k\}$ – множество всех таких наборов. Введем $E(I)=L(e(I))=L(e_{i_1},e_{i_2},\dots,e_{i_m})$ – подпространство в $\mathbf{R}^k$, образованное базисными векторами $e(I)=(e_{i_1},e_{i_2},\dots,e_{i_m})$.

Обозначим через $S_1=\{x\in\mathbf{R}^k \mid \|x\|=1\}$ единичную сферу в $\mathbf{R}^k$.

Пусть $M\subset\mathbf{R}^k$, через $S_1(M)$ обозначим пересечение единичной сферы S_1 с подпространством M, $S_1(M)=S_1\cap M$.

При замене касательной плоскости координатной в методе приведенного градиента предполагается, что на каждом шаге метода выбирается координатная плоскость, наиболее близкая к той касательной, которая реализовалась на данном шаге. Близость понимается в смысле угла между плоскостями ang$(E(I), L(H))$, соответствующее формальное определение угла будет введено далее. Считаем, что касательная плоскость может быть любой (она зависит от конкретной оптимизационной задачи). Каким может оказаться угол между H и ближайшей координатной плоскостью в худшем случае расположения H? Ответ на этот вопрос дается значением следующего максимина:

$$\gamma^* = \max_{H \subseteq H^m(R^k)} \min_{I \in I_k^m} ang(E(I), L(H)),$$

где $H^m(\mathbf{R}^k)$ – множество всех гиперплоскостей в $\mathbf{R}^k$ размерности m. Нам будет удобно искать не значение угла, а значение его косинуса $\cos\gamma^*$.

Значение $\cos\gamma^*$ является общей характеристикой Евклидова пространства, зависит только от m и k, обозначим его через

$$v(k,m) = \min_{H \subseteq H^m(R^k)} \max_{I \in I_k^m} \cos ang(E(I), L(H)). \qquad (1)$$

Цель настоящей работы состоит в получении нетривиальной нижней оценки для $v(k,m)$ в зависимости от произвольных k, m. Таким образом будет построена и нетривиальная верхняя оценка для γ^*.

Определение и свойства угла между гиперплоскостями

Пусть M_1 и M_2 – подпространства в $\mathbf{R}^k$ размерности $m \le k$. Углом между M_1 и M_2 назовем величину

$$ang(M_1, M_2) = \arccos \min_{x \in S_1(M_1)} \left\| \Pr_{M_2} x \right\|$$

где $\Pr_M$ – оператор проектирования в $\mathbf{R}^k$ на M. Аналогичное определение введено в [3] (Глава 12 , стр.520).

Отметим простые свойства угла между подпространствами.

1. Если $m=1$, то ang(M_1, M_2) – не тупой угол между пересекающимися прямыми.
2. Если $m=2$, то ang(M_1, M_2) совпадает с определением не тупого линейного угла двугранного угла между плоскостями.
3. Справедливо равенство ang(M_1, M_2) = ang(M_2, M_1).
4. $\operatorname{ang}(M_1, M_2) = \operatorname{ang}(M_1^\perp, M_2^\perp)$, где знак $\perp$ обозначает ортогональное дополнение в $\mathbf{R}^k$.

Обозначим через $A(M_1, M_2)$ матрицу оператора проектирования $\Pr_{M_2}$ как оператора, действующего из M_1 на M_2.

Пусть $d = (d_1, d_2, ..., d_m)$ – Ортонормированный базис (ОНБ) в M_1, $f = (f_1, f_2, ..., f_m)$ – ОНБ в M_2, тогда $A(M_1, M_2)$ состоит из элементов $a_{ij} = \langle f_i, d_j \rangle$, $A(M_1, M_2) \in \mathbf{R}_{m\times m}$ – матрица оператора проектирования в паре базисов f и d. Здесь и далее $\mathbf{R}_{m\times m}$ – пространство матриц, состоящих из m строк и m

столбцов. Обратим внимание, что $A(M_1,M_2) = G(f,d)$ – матрица Грама пары базисов f и d.

Для вектора $x \in M_1$, заданного своими координатами $\alpha=(\alpha_1, \alpha_2, \ldots, \alpha_m)$ в базисе f ,и вектора $y = \Pr_{M_2} x$, заданного своими координатами $\beta=(\beta_1, \beta_2, \ldots, \beta_m)$ в базисе d ,получим $\beta = A(M_1,M_2)\ \alpha$, т.е. $\beta= G(f,d)\ \alpha$.

Лемма 1. Пусть задано m – мерное подпространство $M \subset \mathbf{R}^k$, $h=(h_1, h_2, \ldots, h_m)$ – ОНБ в M, $E(I) = L(e_{i_1}, e_{i_2}, \ldots, e_{i_m})\ I \in I_k^m$. Тогда $\cos \operatorname{ang}(M,E(I)) = \rho_{min}\ G(h,e(I))$, где ρ_{min} – минимальное сингулярое число матрицы G(h,e(I)).

Доказательство. По определению

$\mathbf{cos}\, ang(M, E(I)) = \min\limits_{x \in S_1(M)} \| \mathbf{Pr}_{E(I)}\, x \|, x \in M$, а так как $h=(h_1, h_2, \ldots, h_m)$ – ОНБ в M, то $x = \sum\limits_{i=1}^{m} \alpha_i h_i$.

Значит $\alpha=(\alpha_1, \alpha_2, \ldots, \alpha_m)^T$ – вектор координат x в разложении по базису h. Знак T обозначает операцию транспонирования. Тогда $\|\mathbf{Pr}_{E(I)}\, x\| = \|G(h,e(I))\alpha\|$. Рассмотрим задачу поиска $\min\limits_{\alpha:\|\alpha\|=1} \| G(h,e(I)) \cdot \alpha \|$, это и будет значение искомого косинуса. Обозначим $G = G(h,e(I\ ))$, $\min\limits_{\alpha:\|\alpha\|=1} \| G(h,e(I)) \cdot \alpha \| = \rho$, тогда $\rho^2 = \min\limits_{\alpha:\|\alpha\|=1} \| G(h,e(I)) \cdot \alpha \|^2$. Имеем

$$\| G(h,e(I)) \cdot \alpha \|^2 = \langle G(h,e(I))\alpha,\ G(h,e(I))\alpha \rangle = \langle\ G^* G\alpha, \alpha\ \rangle.$$

Пусть $\lambda_1, \lambda_2, \ldots, \lambda_m$ – собственные значения оператора G^*G с учетом кратности, $\lambda_i \geq 0$, $i = \overline{1,m}$, а $r_1, r_2, \ldots, r_m$ – собственные векторы. После соответствующей нормировки $r_1, r_2, \ldots, r_m$ образуют ортонормированный базис в $\mathbf{R}^m$. Тогда $\exists\, \mu_i$: $\alpha = \sum\limits_{i=1}^{m} \mu_i r_i$, $G(h,e(I)) \cdot \alpha = \sum\limits_{i=1}^{m} \mu_i r_i \lambda_i$, $\| \alpha \|^2 = \left(\sum\limits_{i=1}^{m} \mu_i r_i \right)^2$,

и так как $\|r_i\| = 1$, то $\left(\sum\limits_{i=1}^{m} \mu_i r_i \right)^2 = \sum\limits_{i=1}^{m} \mu_i^2$.

Таким образом

$$\min_{\alpha:\|\alpha\|=1} \langle\ G^* G\alpha, \alpha\ \rangle = \min_{\mu:\sum_{i=1}^{m}\mu_i^2=1} \langle\ \sum_{i=1}^{m} \mu_i r_i \lambda_i, \sum_{i=1}^{m} \mu_i r_i\ \rangle = \min_{\mu:\sum_{i=1}^{m}\mu_i^2=1} \sum_{i=1}^{m}\ \lambda_i \mu_i^2 = \min_{i=\overline{1,m}} \lambda_i.$$

Докажем последнее равенство. Расположим λ_i в порядке возрастания, при этом каждое из собственных чисел запишется столько раз, какова его кратность. Пусть $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$, тогда

$$\lambda_m(\mu_1^2 + \mu_2^2 + \ldots + \mu_m^2) \leq \lambda_1\mu_1^2 + \lambda_2\mu_2^2 + \ldots + \lambda_m\mu_m^2 \leq \lambda_1(\mu_1^2 + \mu_2^2 + \ldots + \mu_m^2),$$

причем знаки равенства достигаются в случаях ($\mu_m = 1$, $\mu_1 = \mu_2 = ... = \mu_{m-1} = 0$) и ($\mu_1 = 1$, $\mu_2 = \mu_3 = ... = \mu_m = 0$), соответственно. В результате $\rho^2 = \min\limits_{i=\overline{1,m}} \lambda_i$, где λ_i – собственные числа матрицы G^*G, а искомый минимум равен $\rho = \min\limits_{i=\overline{1,m}} \rho_i$, где ρ_i – сингулярные числа матрицы G, ибо $\rho_i^2 = \lambda_i, i = \overline{1,m}$. Следовательно $\rho = \rho_{\min}$. Лемма доказана.

Получение оценки.

Введем обозначения для миноров. Минор матрицы D, расположенный в строках с номерами $l_1, l_2, ..., l_p$ и столбцах с номерами $j_1, j_2, ..., j_p$, будем обозначать $D^{l_1, l_2, ..., l_p}_{j_1, j_2, ..., j_p}$.

Далее для получения оценки будет использоваться следующая формула. Формула Бине-Коши [4] (Глава 1, стр.41).

Пусть даны матрицы: $A \in \mathbf{R}_{m\times k}$, $B \in \mathbf{R}_{k\times m}$ и $C = A \cdot B \in \mathbf{R}_{m\times m}$. Тогда $\det C = \sum\limits_{0 \le s_1 < s_2 < ... < s_m} \det A^{1,2,...,m}_{s_1, s_2, ..., s_m} \det B^{s_1, s_2, ..., s_m}_{1,2,...,m}$.

Теорема. Для любого подпространства $M \subset \mathbf{R}^k$: $\dim M = m \le k$ найдется такое координатное подпространство $I^* \in I^m_k$, что $\operatorname{ang}(M, E(I^*)) \le \arccos\left(\dfrac{1}{C^m_k}\right)^{1/2}$, где через C^m_k обозначено число сочетаний из k по m.

Доказательство. Пусть $h = \{h_1, h_2, ..., h_m\}$ – ОНБ в M, матрица $H = \langle h_i, e_j \rangle \in \mathbf{R}_{m\times k}$. По Лемме 1 $\forall I \in I^m_k$ $\cos \operatorname{ang}(M, E(I)) = \rho_{min}\, G(h, e(I))$. Тогда по формуле Бине-Коши

$$\det H^* H = \sum_{I \in I^m_k} \det G(h, e(I\ ))^2 .$$

Так как $H^* H = G(h,h)$ и $\det (H^T H) = 1$, то найдется $I^* \subset I^m_k$, такое, что

$$\det G(h, e(I^*)) \ge \sqrt{\frac{\det(H^T H)}{C^m_k}} = \sqrt{\frac{1}{C^m_k}}.$$

Обозначим $G = G(h, e(I^*))$. Пусть ρ_1, $\rho_2, ..., \rho_m$ – сингулярные числа $G(h, e(I^*)) = G$, $\lambda_1, \lambda_2, ..., \lambda_m$ – собственные значения G^*G, тогда $\rho_i^2 = \lambda_i$, $i = \overline{1,m}$. Так как $\det G^* = \det G$, где G^* – матрица сопряженная к матрице G, то

$$|\det G|^2 = |\det G^* G| = \prod_{i=1}^m \rho_i^2 = \left(\prod_{i=1}^m \rho_i\right)^2 .$$

Тогда

$$\left|\det G(h,e(I^*))\right| = \prod_{i=1}^{m}\rho_i = \left(\prod_{i=1}^{m}\lambda_i\right)^{\frac{1}{2}},$$

т.е. $\left|\det G(h,e(I^*))\right|^2 = \lambda_1\cdot\lambda_2\cdot\ldots\cdot\lambda_m \le \lambda_{\min}\cdot\lambda_{\max}^{m-1}$,где $\lambda_{\min}$ $и$ $\lambda_{\max}$ –минимальное и максимальное собственные значения. Значит $\lambda_{\min} \ge \left(\frac{1}{C_k^m}\right)^{1/2}\lambda_{\max}^{m-1}$.

Рассмотрим матрицу оператора $G^*G \in \mathbf{R}_{\mathbf{m\times m}}$ и ее максимальное собственное значение $\lambda_{\max}$, соответствующий собственный вектор обозначим через p , $Ap = \lambda_{\max}p$, $\|p\|=1$ [5] (Глава 15, стр.240).

Для евклидовой нормы оператора $\|G^*Gp\| = |\lambda_{\max}|\|p\|$ [5] (Глава 18 , стр.307) и $\|G^*Gp\| \le \|G^*G\|\|p\|$, откуда следует , что $|\lambda_{\max}| \le \|G^*G\|$. Заметим, что $\|G^*G\| \le \|G^*\|\|G\|$, и так как $G^* = G$, то $\|G^*G\| \le \|G^2\|$. Значит, $|\lambda_{\max}| \le \|G^*G\| \le \|G^2\|$. Докажем , что норма оператора $G^*G \le 1$. По определению $\|G^*G\| = \max\limits_{p:\|p\|=1}\frac{\|G^*Gp\|}{\|p\|} = \max\|G^*Gp\|$. Так как G^*G – оператор проектирования , то $p = G^*Gp + \left(G^*Gp\right)^{\perp}$,где G^*Gp – ортогональная проекция, а$(G^*Gp)^{\perp}$ – ортогональная составляющая вектора p. Тогда $\|p\|^2 = \|G^*Gp\|^2 + \left\|\left(G^*Gp\right)^{\perp}\right\|^2$, т.е. $\|G^*Gp\| \le \|p\| = 1$. Таким образом $|\lambda_{\max}| \le \|G^*G\| \le 1$. В результате $\cos ang(M,e(I^*)) = \rho_{\min}G(h,e(I^*)) \ge \left(\frac{1}{C_k^m}\right)^{\frac{1}{2}}$, что и требовалось доказать.

Литература.

1. Полак Э. Численные методы оптимизации – М., Мир, 1974.
2. Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации – М., Наука, 1986.
3. Голуб Дж., Ван Лоун Ч. Матричные вычисления – М., Мир, 1999.
4. Ланкастер П. Теория матриц. – М., Наука, 1978.
5. Ильин В.А., Ким Г.Д. Линейная алгебра и аналитическая геометрия. Издательство Московского Университета , 1998.