PROJECTION METHODS FOR LINEAR OPTIMIZATION

P. Kosmol

Известия Национальной Академии Наук Армении. Математика, том 36, № 6, 2001

A method of successive projections for the solution of LP-problems is presented, where the computation of the optimal solution is replaced by determination of fixed points of nonexpanding operators. The non-differentiable projection onto the positive cone is then replaced by a sequence of smooth approximations that facilitates the application of rapidly convergent numerical methods. As the corresponding sequence of operators turns out to be equicontinuous, stability of the sequence of solutions follows. The method can be highly efficient, in particular for semi-infinite type problems. A particular feature of the method is that it does not require presence of interior points in the restriction set (as Karmarkar's method does).

§1. INTRODUCTION

The method described below can be applied for numerical solution of a number of approximation and optimization problems like Chebyshev-approximation, L^1-approximation, One-sided L^1-approximation, Markov moment problems, Semi-infinite linear optimization.

However, we present an algorithm tailored for linear programming problems. It uses the method of successive projections analogous to the Kaczmarz / v. Neumann method.

For the LP–problem we use the following notation :

Let $c, x \in \mathbb{R}^n$, $A = \begin{pmatrix} a_1 \\ . \\ . \\ . \\ a_m \end{pmatrix} \in L(\mathbb{R}^n, \mathbb{R}^m)$, $b \in \begin{pmatrix} b_1 \\ . \\ . \\ . \\ b_m \end{pmatrix} \in \mathbb{R}^m$. Then we consider the following problem :

$$\min\{\langle c, x\rangle x \mid A\,x = b, \quad x \geq 0\}.$$

§2. THE PROJECTION METHOD

Method 2.1. We chose the mapping $P = P_m \circ P_{m-1} \circ ... \circ P_1$ of the successive projections P_i onto the hyperplanes

$$S_i = \{s \in \mathbb{R}^n \mid < a_i, s >= b_i\} \quad \text{for} \quad i \in \{1, ..., m\},$$

that is

$$P_i(x) = x + \frac{(b_i - < a_i, x >)}{\|a_i\|^2} a_i$$

and the projection P_K onto the positive cone $\mathbb{R}^n_{\geq 0}$, given by

$$P_K(x) := ((x_1)_+, ..., (x_n)_+)).$$

Furthermore we introduce a regularization parameter $\alpha > 0$ to determine fixed points of

$$T_\alpha(x) := P_K \circ P(x) + \alpha \cdot c.$$

For a positive sequence $(\alpha_n)_{n\in\mathbb{N}}$ tending to zero the sequence $(x_n)_{n\in\mathbb{N}}$ of fixed points of nonexpanding mappings $T_{\alpha_n}(x)$ every limit point of $(x_n)_{n\in\mathbb{N}}$ is a solution of the LP-problem.

The above method evolved in the context of a general theory on **two-stage** solutions of variational inequalities. For the determination of the fixed points the usual Picard iteration was used. The disadvantage of the above method is the nondifferentiability of the projection P_K on $\mathbb{R}_{\geq 0}$, which prevents the application of rapidly convergent numerical methods (Newton like methods), where the fixed points are determined via solution of the corresponding nonlinear equation.

The modification of the "old" method now consists in a smoothing of the projection P_K, more specifically : the function $s \mapsto (s)_+$ is replaced by a one-dimensional smoothing $\varphi_\alpha : \mathbb{R} \to \mathbb{R}$ that approximates the $(\cdot)_+$-function.

The projection P_K is then replaced by $P_\alpha = (\varphi_\alpha(x_1), ..., \varphi_\alpha(x_n))$. By use of the Newton-method the nonlinear equation $F_\alpha := P_\alpha \circ P(x) + \alpha c - x = 0$ is solved. The directional derivation of F_α turns out to be $F'_\alpha(x, z) = P'_\alpha(P(x)) \cdot P_0(z)$. The derivative of P_α is then given by the diagonal matrix

$$P'_\alpha(x) = \begin{pmatrix} \varphi_\alpha(x_1) & 0 \\ 0 & \varphi'_\alpha(x_n) \end{pmatrix}$$

and P_0 is the linear part of the affine mapping P, which corresponds to the successive projection onto the hyperplanes $\{x | \langle a_i, x \rangle = 0\}$, $i \in \{1, ..., n\}$.

Theorem 2.2. Let X be a normed space, $A: X \to X^*$ and for $(A_n : X \to X^*)$ let $L = \overline{\lim}_{n\to\infty}\{x|A_n x = 0\} \subset \{x|Ax = 0\} =: S$. Let further (a_n) be a sequence of positive numbers such that $a_n(A_n - A)$ converges continuously to a mapping $D: X \to X^*$. Let $\overline{x} \in L$, then for all $x \in S$ the following inequality $\langle D\overline{x}, x - \overline{x}\rangle \geq 0$ holds.

Proof : Let $x_k \epsilon \{x|A_k x = 0\}$ such that (x_k) converges to an $\overline{x} \in X$, i.e. $\overline{x} \in L$. Let $x \in S$, i.e. $Ax = 0$. Since A monotone and radially continuous it follows from Vainberg [6] :

$$\alpha_k\langle (A_k - A)x_k, x - x_k\rangle \geq 0$$

Since $\alpha_k(A_k - A)$ converges to D it follows that $\alpha_k(A_k - A)x_k$ converges to $D\overline{x}$ in the norm.

Application of this theorem to the "old" method yields for $A = P_K \circ P - I$ and the sequence $A_n = P_K \circ P + \alpha_n c - I$ for $a_n = \frac{1}{\alpha_n}$ that $a_n(A_n - A) = c$, which is a constant sequence obviously continuously convergent to the constant function $D \equiv c$.

If $\overline{x}$ is a limit point of the sequence of the fixed points of the sequence of operators (A_n), then for all $x \in S$ we obtain $\langle c, x - \overline{x}\rangle \geq 0$. Note that $\overline{x}$ is a solution of the original LP-problem, once we have shown that $L \subset S$.

Application of the above theorem to the "smoothed method" yields for $A_n = F_{\alpha_n}$ and $A = P_K \circ P - I$ a condition for φ_{α_n} that enforces continuous convergence. We have for $a_n = \frac{1}{\alpha_n}$

$$a_n(A_n - A) = a_n(P_{\alpha_n} \circ P + \alpha_n c - P_k \circ P) = \frac{1}{\alpha_n}(P_{\alpha_n} \circ P - P_k \circ P) + c.$$

It follows that, if $\frac{1}{\alpha_n}(\varphi_{\alpha_n} - (\cdot)_+) \to 0$ uniformly on compact subsets of $\mathbb{R}$, then $a_n(A_n - A)$ converges continuously to c.

Remark. The above method can be extended to convex functionals f if in the operator A_n the vector c is replaced by the derivative f'. Then $a_n(A_n - A)$ converges continuously to $f' = D$. The characterization theorem of convex optimization yields that every point of accumulation is a minimal solution of f on S. The computational part of the method remains essentially unchanged.

Example 2.3. For $\beta = \beta(\alpha)$ let

$$\varphi_\alpha(s) = \begin{cases} s & \text{for} \quad s \geq \beta \\ (s+\beta)^2/4\beta & \text{for} \quad -\beta \leq s \leq \beta \\ 0 & \text{for} \quad s \leq -\beta, \end{cases}$$

hence $\varphi_\alpha - (\cdot)_+ \equiv 0$ on $(-\infty,-\beta)\cup(\beta,\infty)$ and on $(-\beta,\beta)$ we obtain

$$|\varphi_\alpha(s) - (s)_+| = \left|\frac{1}{\alpha}\left(\frac{(s+\beta)^2}{4\beta} - (s)_+\right)\right| \leq \frac{1}{\alpha}\cdot\frac{\beta}{4}.$$

If $\frac{\beta_n}{\alpha_n} \to_{n\to\infty} 0$, then $\frac{1}{\alpha_n}(\varphi_{\alpha_n} - (\cdot)_+) \to 0$ uniformly on compact subsets of $\mathbb{R}$ and hence $a_n(A_n - A)$ converges continuously to c.

Theorem 2.4. Let for all $s \in \mathbb{R}$ and all $\alpha > 0$, $|\varphi'_\alpha(s)| \leq \gamma < 1$, then $T_\alpha = P_\alpha \circ P + \alpha c$ is a contraction and $F_\alpha = I - T_\alpha$ is strongly monotone (well posed problem). Moreover $F'_\alpha(x)$ is positive definite everywhere.

Proof : We first show that T_α is a contraction. Using the mean value theorem and the nonexpandivity of P it follows

$$\|T_\alpha(x) - T_\alpha(y)\| \leq \|P_\alpha(Px) - P_\alpha(Py)\| =$$
$$= \left\|\int_0^1 (P'_\alpha(Py - t(Px - Py)(Py - Px)\,dt\right\| =$$
$$= \int_0^1 \|P'_\alpha(Py - t(Px - Py)(Py - Px)\|\,dt \leq \gamma\|Py - Px\| \leq \|x - y\|$$

Thus, T_α is a contraction and hence $F_\alpha = I - T_\alpha$ strongly monotone, because

$$\langle (I - T_\alpha)x - (I - T_\alpha)y, x - y\rangle = \|x - y\|^2 - \langle (T_\alpha x - T_\alpha y, x - y\rangle \geq$$
$$\geq \|x - y\|^2 - \|T_\alpha x - T_\alpha y\|\|x - y| \geq (1-\gamma)\|x - y\|^2.$$

Thus it follows for all $x, v \in \mathbb{R}^n$ $\langle F_\alpha(x + tv) - F_\alpha(x), tv\rangle \geq (1-\gamma)t^2\|v\|^2$ and for

$$t \mapsto \varphi(t) := \langle F_\alpha(x + tv) - F_\alpha(x), v\rangle \geq (1-\gamma)t\|v\|^2,$$

hence for $v \neq 0$:

$$\varphi'(0) = \langle F'_\alpha(x)v, v\rangle \geq (1-\gamma)\|v\|^2 > 0.$$

Example 2.5. Let

$$\varphi'_\alpha(s) = \begin{cases} 0 & s \leq 0 \\ \dfrac{1-\beta}{\beta}s & 0 \leq s \leq \beta \\ (1-\beta) & s > \beta \end{cases}, \quad \varphi_\alpha(s) = \begin{cases} 0 & s < 0 \\ \dfrac{1-\beta}{2\beta}s^2 & 0 \leq s \leq \beta \\ (1-\beta)(s - \frac{\beta}{2}) & s > \beta, \end{cases}$$

then for $0 \leq s \leq \beta$ we obtain

$$|\frac{1}{\alpha_n}(\varphi_{\alpha_n}(s) - (s)_+)| = \left|\frac{1}{\alpha}\frac{(1-\beta)s^2}{2\beta} - s\right| = \frac{s}{\alpha}\left|\frac{(1-\beta)s - 2\beta}{\beta}\right| \leq \frac{\beta}{\alpha}\cdot 3 \to_{\frac{\beta}{\alpha}\to 0} 0$$

and for $s > \beta$

$$|\frac{1}{\alpha_n}(\varphi_{\alpha_n}(s) - (s)_+)| = \left|\frac{1}{\alpha}((1-\beta)(s - \frac{\beta}{2}) - s\right| =$$
$$= \frac{1}{\alpha}\left|-\beta s - (1-\beta)\frac{\beta}{2}\right| = \frac{\beta}{\alpha}\left|\left(s + \frac{(1-\beta)}{2}\right)\right| \to 0.$$

The following lemma expresses the fact that the angle between $x - k_0$ and $k - k_0$ is obtuse :

Lemma 2.6. Let $K \subset \mathbb{R}^n$ convex and let k_0 be best Euclidean approximation of $x \in \mathbb{R}^n$ (with respect to K). Then for all $k \in K$ the following inequality holds :

$$||x - k||^2 \geq ||x - k_0||^2 + ||k_0 - k||^2.$$

The following theorem states that a composition of projections is a contraction with respect to its fixed points :

Theorem 2.7. Let S be nonempty and bounded. Then for every $x_0 \in S$ there is a $r > 0$ and a $\gamma \in (0,1)$ such that for all $x \in \mathbb{R}^n$ with $||x - x_0|| > r$ we have

$$||x_0 - P(x)|| \leq \gamma ||x_0 - x||.$$

Proof: Let $x_0 \in S$, and for $i \in \{1, ..., n\}$ let $a_{m+i} := -e_i$ and $b_{m+i} = 0$. Furthermore, let $M > 0$ be chosen such that $K(0, M/2) \supset S$. Then for all $x \notin K(0, M)$

$$\beta(x) := \max\left\{\frac{|\langle x - x_0, a_i\rangle|}{||a_i||} | 1 \leq i \leq m+n\right\} > 0$$

and using the compactness of the sphere $S(0, M)$ it follows :

$$\gamma := \min\{\beta(x) | x \in S(0, M)\} > 0.$$

Let

$$\delta := \max\{\frac{b_i - \langle a_i, x_0\rangle|}{||a_i||} | 1 \leq i \leq m+n\}.$$

As $x_0 \in S$ obviously we have $\delta = ||x_0||_\infty \geq 0$. Chose $r > \frac{2\delta M}{\gamma}$ and $x \notin K(0, r)$. For some $j \in \{1, ..., m+n\}$ we have

$$\frac{M\langle x - x_0, a_j\rangle}{||x - x_0|| ||a_j||} = \max\{\frac{M\langle x - x_0, a_i\rangle}{||x - x_0|| ||a_i||} | 1 \leq i \leq m+n\}.$$

It follows that

$$\frac{|b_j - \langle x, a_j\rangle|}{||a_j||} = \frac{|b_j - \langle x_0, a_j\rangle - \frac{1}{M}||x - x_0||\langle \frac{M(x-x_0)}{||x-x_0||}, a_j\rangle|}{||a_j||} \geq$$
$$\geq \frac{\gamma}{M}||x - x_0|| - \delta \geq \frac{\gamma}{2M}||x - x_0||.$$

∎

Let $\widetilde{P}_0(x) := x$ and for $k \in \{1, ..., m+n\}$: $\widetilde{P}_k := P_k \circ ... \circ P_1$. Because of $P_j(x) \in S_j$ it follows that $||\widetilde{P}_j(x) - x|| \geq d(x, S_j)$ and by using the triangle inequality

$$\sum_{k=1}^{j} ||\widetilde{P}_k(x) - \widetilde{P}_{k-1}(x)|| \geq ||\widetilde{P}_j(x) - x|| \geq \frac{\gamma}{2M}||x - x_0||$$

Then there is $l \in \{1,...,j\}$ such that

$$\|\widetilde{P}_l(x) - \widetilde{P}_{l-1}(x)\| \geq \frac{\gamma}{2mM}\|x - x_0\|$$

By Lemma 2.6 we obtain using the nonexpandivity of $\widetilde{P}_l$:

$$\|x - x_0\|^2 \geq \|\tilde{P}_{l-1}(x) - P_{l-1}(x_0)\|^2 \|\widetilde{P}_{l-1}(x) - x_0\|^2 \geq \|\widetilde{P}_{l-1}(x) - P_{l-1}(x)\|^2 +$$
$$+ \|\widetilde{P}_l(x) - x_0\|^2 \geq \frac{\gamma^2}{4m^2M^2}\|x - x_0\| + \|\widetilde{P}_l(x) - x_0\|^2.$$

For $c := \sqrt{1 - (\frac{\gamma}{2mM})^2}$ using once more the nonexpandivity of projections we establish the claim of the theorem.

We still have to answer the stability question $\emptyset \neq \overline{\lim}\{x|A_n x = 0\} \subset \{x|Ax = 0\}$. The following theorem gives a positive answer for examples 2.3 and 2.5 :

Theorem 2.8. Let $(B_k : \mathbb{R}^n \to \mathbb{R}^n)$ be a sequence of nonexpanding operators that converges pointwise to $P_K \circ P$. Then there is $N \in \mathbb{N}$ such that for $k \geq N$ the set of fixed points $Fix(B_k) \neq \emptyset$ and $\bigcup_{k=N}^{\infty} Fix(B_k)$ is bounded. Moreover, every point of accumulation of a sequence $(x_k \in Fix(B_k))$ is a fixed point of $P_K \circ P$.

Proof : Apparently, the sequence (B_k), being a sequence of nonexpanding operators, is equicontinuous. Therefore, pointwise convergence of B_k to $P_K \circ P$ implies continuous convergence, which is equivalent to uniform convergence on compact subsets of $\mathbb{R}^n$. We first observe : let $x_0 \in Fix(P_K \circ P)$ then

$$\|P_K \circ P(x) - x_0\| \leq \|x - x_0\|,$$

i.e. any ball with center x_0 is mapped into itself via $P_K \circ P$. In particular, the interior of such a ball is mapped into the interior. Let $R > r$ we have by Theorem 2.7 for $x \in S(x_0, R)$ that $P_K \circ P(x)$ is an interior point of $K(x_0, R)$. For reasons of compactness there is $\varepsilon > 0$ such that

$$P_K \circ P(K(x_0, R)) \subset K(x_0, R - \varepsilon).$$

Because of the uniform convergence of (B_k) there is a $K \in \mathbb{N}$ such that $B_k(K(x_0, R) \subset K(x_0, R - \frac{\varepsilon}{2})$. Because of the nonexpandivity of B_k we have for $k \geq N$

$$\emptyset \neq Fix(B_k) \subset K(x_0, R).$$

Continuous convergence guarantees that every point of accumulation of fixed points of B_k is a fixed point of $P_K \circ P$.

We summarize the results for the second example :

1. The Newton-method converges superlinearly (because of the regularity of the Jacobian).

2. $F_{\alpha_n}(x) = 0$ is uniquely solvable and the sequence of the solution is bounded.

3. The accumulation point is a solution of the original LP-problem.

The above algorithms have been successfully tested. The method turned out to be particularly effective for many variables and a moderate number of restrictions, among those : discretizations of moment problems, and the dual of semi-infinite linear optimization problems.

A more general framework for treating stability questions is given by the following scheme [4] :

Theorem 2.9 (K. Müller-Wichards). Let $A_n : \mathbb{R}^n \to \mathbb{R}^n$ be a sequence of continuous operators that converges uniformly on compact subsets of $\mathbb{R}^n$ to an operator $A : \mathbb{R}^n \to \mathbb{R}^n$ with the property :

there exists a ball $K(x_0, r), r > 0$ such that $\langle Ax, x - x_0 \rangle > 0$ for all x in the sphere $S(x_0, r)$.

Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the equations $A_n x = 0$ have a solution in $K(x_0, r)$. Furthermore every point of accumulation of these solutions is a solution of $Ax = 0$.

The above theorem is a consequence of the following well known lemma.

Lemma 2.10. Let $f : \mathbb{R}^n \to \mathbb{R}^n$. If there is $x_0 \in \mathbb{R}^n$ and $R > 0$ such that $\langle f(x), x - x_0 \rangle \geq 0$ for all $x \in S(x_0, R)$, then the nonlinear equation $f(x) = 0$ has a solution in $\overline{K}(x_0, R)$.

Proof : Otherwise Brouwer's fixed point theorem applied to the mapping

$$x \mapsto g(x) := -R\left(\frac{f(x)}{\|f(x)\|}\right) + x_0$$

would lead to a contradiction.

In particular, the above principle can be applied to the limiting function $P_K \circ P$ in the following sense : by Theorem 2.7 we have for all $x \in S(x_0, R)$:

$$\langle (I - P_K \circ P)x, x - x_0 \rangle = \langle (I - P_K \circ P)x - (I - P_K \circ P)x_0, x - x_0 \rangle \geq$$

$$\geq \|x - x_0\|^2 - \|P_K \circ Px - x_0\| \|x - x_0\| = (1 - \gamma)\|x - x_0\|^2 = (1 - \gamma)\, R^2.$$

A large class of operators can be treated using the above stability principle, where pointwise convergence already implies uniform convergence on compact subsets. Among them are : convex operators, monotone operators, component-wise convex operators, compositions conserving equicontinuity.

АБСТРАКТ. Представлен метод последовательных проекций для решения LP–задач, в котором вычисление оптимального решения заменяется определением фиксированных точек нерастягивающихся операторов. Недифференцируемая проекция на положительный конус заменяется на последовательность гладких приближений, что облегчает использование быстро сходящихся численных методов. Так как соответствующая последовательность операторов оказывается экви-непрерывна, то последовательность решений является стабильной. Этот метод может быть очень эффективным, в частности, для задач полубесконечного типа. Преимущество этого метода также состоит в том, что он не требует наличия внутренних точек в суженном множестве (как это делается в случае метода Кармаркара).

REFERENCES

1. P. Kosmol, Optimierung und Approximation, de Gruyter Lehrbuch, Berlin, New York, 1991.
2. P. Kosmol, Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben, B.G. Teubner Studienbücher, Stuttgart, zweite Auflage, 1993.
3. P. Kosmol, D. Müller-Wichards, Optimierung in Orlicz-Räumen (Monography, in preparation).
4. P. Kosmol, D. Müller-Wichards, "On Stability for Families of nonlinear Equations (in preparation).
5. Kosmol, P. : Ein Algorithmus für Variationsungleichungen und lineare Optimierungsaufgaben, Deutsch - französiches Treffen zur Optimierungstheorie, Hamburg, 1986.
6. M. M. Vainberg, Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations, Wiley and Sons, 1973.

21 September 2001

University of Kiel,
Germany