UN ESTUDIO SOBRE LA APLICACIÓN DEL MÉTODO DEL GRADIENTE ESPECTRAL PROYECTADO PARA EL PROBLEMA MULTIOBJETIVO

Nadia Soledad Fazzio[†‡] y María Laura Schuverdt[†‡]

[†]*Departamento de Matemática, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina.*
[‡]*CONICET, Argentina.*

Resumen: En este trabajo consideramos el problema de optimización multiobjetivo con restricciones de igualdad y desigualdad. Las soluciones de estos problemas son los conocidos puntos Pareto óptimos o Pareto óptimos débiles pero en este trabajo estamos interesados en puntos estacionarios, es decir, puntos que cumplen una condición necesaria de optimalidad.
Basados en los resultados obtenidos para el método del gradiente espectral proyectado para el caso escalar proponemos alternativas para definir un método del mismo para resolver el problema multiobjetivo. Estudiamos posibilidades para definir coeficientes espectrales adecuados que permitan demostrar convergencia global a puntos estacionarios del problema multiobjetivo. En este trabajo la dirección de búsqueda lineal utiliza el coeficiente espectral y la búsqueda lineal es monótona.

Palabras clave: *OPTIMIZACIÓN MULTIOBJETIVO, GRADIENTE ESPECTRAL PROYECTADO, BÚSQUEDA LINEAL MONOTONA*
2000 AMS Subject Classification: 49M37 - 90C29

1. INTRODUCCIÓN

Estamos interesados en el siguiente problema de optimización multiobjetivo con restricciones:

$$\begin{array}{ll} minimizar & F(x) \\ sujeto\ \ a & x \in C \end{array} \tag{1}$$

donde $F : \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (F_1(x), \cdots, F_m(x))$ es continuamente diferenciable y $C \subset \mathbb{R}^n$ es cerrado y convexo. Para este problema, buscamos un Pareto óptimo (o punto eficiente), es decir, un punto factible $(x^* \in C)$ de tal manera que no existe $x \in C$ con $F(x) \leq F(x^*)$ y $F(x) \neq F(x^*)$; o un Pareto óptimo débil (o punto débilmente eficiente), es decir, un punto factible x^* tal que no existe $x \in C$ con $F(x) < F(x^*)$.

Para el problema multiobjetivo con restricciones, una condición necesaria de optimalidad para un punto $x^* \in C$ es ser estacionario (o crítico) es decir,

$$JF(x^*)(C - \{x^*\}) \cap [-\mathbb{R}^m_{++}] = \emptyset$$

donde

$$JF(x^*)(C - \{x^*\}) := \{JF(x^*)(x - x^*) : x \in C\}.$$

Es decir, $x^* \in C$ es un punto estacionario del problema (1) si, y sólo si, para todo $x \in C$,

$$JF(x^*)(x - x^*) \notin -\mathbb{R}^m_{++},$$

es decir $JF(x^*)(x - x^*)$ no es menor que 0.

Luego esta condición es equivalente a: $x^* \in C$ es un punto estacionario para F si, y sólo si, para todo $x \in C$,

$$\max_{i=1,\ldots,m} \nabla F_i(x^*)^T (x - x^*) \geq 0. \tag{2}$$

Observemos que para el problema de optimización escalar con restricciones $(m = 1)$ la condición (2) se reduce a la condición clásica de ser estacionario $\langle \nabla F(x^*), x - x^* \rangle \geq 0$ para todo $x \in C$.

Luego, un punto $x^* \in C$ no es estacionario si existe una dirección $v \in C - \{x^*\}$ tal que $JF(x^*)v < 0$. El siguiente resultado muestra que v es una dirección de descenso para F en x (es decir, existe $\epsilon > 0$ tal que $F(x + tv) < F(x)$ para todo $t \in (0, \epsilon]$).Este resultado nos permite implementar una regla tipo Armijo para los métodos de descenso multiobjetivo.

Lema 1 *Sea $F : \mathbb{R}^n \to \mathbb{R}^m$ continuamente diferenciable, $v \in C - \{x\}$ tal que $JF(x)v < 0$, es decir, $\nabla F_i(x)v < 0$ para $i = 1, \cdots, m$ y $\sigma \in (0, 1)$. Entonces existe algún $\epsilon > 0$ (que puede depender de x, v y σ) de tal manera que*

$$F(x + tv) < F(x) + \sigma t JF(x)v$$

para cualquier $t \in (0, \epsilon]$

En particular, v es una dirección de descenso para F en x.

Se puede ver la demostración en [3].
En [3] se muestra como se extiende la dirección de búsqueda, originalmente definida para $C = \mathbb{R}^n$, a C convexo y cerrado.
Originalmente, para $C = \mathbb{R}^n$ se define

$$v(x) = \underset{v}{argmin} \quad \varphi_x(v) + \tfrac{\|v\|^2}{2},$$

luego, para C convexo y cerrado, se define como

$$v(x) := \underset{v \in C - \{x\}}{argmin} \quad \beta\varphi_x(v) + \tfrac{\|v\|^2}{2} \tag{3}$$

donde $\beta > 0$ es un parámetro y $\varphi_x(v) = \max_i \nabla F_i(x)^T v$.

Observar que, por la convexidad la dirección $v(x)$ está siempre bien definida. Ahora se define el valor funcional optimo $\theta(.)$ para (3) como

$$\theta(x) := \min_{v \in C - \{x\}} \beta\varphi_x(v) + \frac{\|v(x)\|^2}{2} = \beta\varphi_x(v(x)) + \frac{\|v(x)\|^2}{2} \tag{4}$$

Se puede demostrar el siguiente teorema:

Teorema 1 *Sea $v : C \to \mathbb{R}^n$ y $\theta : C \to \mathbb{R}$ dado por*

$$v(x) := \underset{v \in C - \{x\}}{argmin} \quad \beta\varphi_x(v) + \tfrac{\|v\|^2}{2}$$

y

$$\theta(x) := \beta\varphi_x(v(x)) + \frac{\|v(x)\|^2}{2}.$$

Entonces las siguientes afirmaciones valen.

1. *$\theta(x) \leq 0$ para todo $x \in C$.*
2. *La función $\theta(.)$ es continua.*
3. *Las siguientes condiciones son equivalentes:*
    a) *El punto $x \in C$ no es estacionario.*
    b) *$\theta(x) < 0$.*
    c) *$v(x) \neq 0$*

En particular, x es estacionario si y slo si $\theta(x) = 0$.

Se puede ver la demostración en [3].
De igual manera sabemos que continua valiendo lo mencionado si consideramos

$$v(x) := \underset{v \in C - \{x\}}{argmin} \quad \varphi_x(v) + \tfrac{\|v\|^2}{2\beta}$$

y

$$\theta(x) := \varphi_x(v(x)) + \frac{\|v(x)\|^2}{2\beta}, \text{ para } \beta > 0.$$

Nota 1 *En este trabajo, basados en los resultados obtenidos para el método del gradiente espectral proyectado para el caso escalar [1], consideramos una sucesión $\{\beta_k\}$ que varia de un paso al siguiente del algoritmo, que es acotada y que posee propiedades similares a las que se tienen en el caso escalar.*

2. Algoritmo de gradiente espectral proyectado para el problema multiobjetivo y análisis de convergencia.

Ahora estamos en condiciones de definir una extensión del método de gradiente espectral proyectado del problema escalar para el problema de optimización multiobjetivo con restricciones (1).
Algoritmo Elegir parámetros fijos $\beta_0 > 0$, $\sigma \in (0, 1)$ $0 < \beta_{min} < \beta_{max} < \infty$ y $x^0 \in C$. Tomar $k = 0$,

1. Calcular la dirección de búsqueda. Calcular

$$v_k = v(x_k) := \underset{v \in C - \{x_k\}}{argmin} \quad \varphi_{x_k}(v) + \tfrac{\|v\|^2}{2\beta_k}$$

   donde $\varphi_{x_k}(v) := \max_{i=1,\ldots,m} \nabla F_i(x_k)^T v$.

2. Chequear optimalidad. Calcular $\theta(x_k) = \varphi_{x_k}(v_k) + \frac{1}{2\beta_k}\|v_k\|^2$. Si $\theta(x_k) = 0$, entonces parar.

3. Calcular la longitud de paso. Eligir t_k como el mayor $t \in \left\{\frac{1}{2^j} : j = 0, 1, 2, ...\right\}$ tal que

$$F(x_k + tv_k) \leq F(x_k) + \sigma t JF(x_k)v_k.$$

4. Tomar $x_{k+1} = x_k + t_k v_k$. Definir un coeficiente β_{k+1} tal que $\beta_{min} < \beta_{k+1} < \beta_{max}$, $k \leftarrow k + 1$, e ir al paso 2.

Nota 2 *En el caso escalar [1] el coeficiente espectral se define como*

$$\beta_{k+1} = \begin{cases} \min\{\beta_{max}, \max\{\beta_{min}, \frac{s_k^T s_k}{s_k^T y_k}\}\} & si \quad s_k^T y_k > 0 \\ \beta_{max} & sino \end{cases} \tag{5}$$

Se sabe que, si $\nabla f(x_k)^T v \geq 0\ \forall v \in \mathbb{R}^n$ entonces $\nabla f(x_k) = 0$ (es decir x_k es un punto estacionario). Para el caso multiobjetivo si $\max_{i=1,\ldots,m} \nabla F_i(x_k)^T v \geq 0$ entonces x_k es un punto estacionario. Luego en el caso escalar $s_k^T y_k = s_k^T(\nabla f(x_{k+1}) - \nabla f(x_k))$ es decir, $s_k^T y_k$ se define a partir de la medida dada por $\nabla f(x)^T v$ evaluado en x_k y en x_{k+1}. En el caso multiobjetivo esa medida está dada por $\varphi_{x_k}(v) := \max_{i=1,\ldots,m} \nabla F_i(x_k)^T v$. Por eso podríamos considerar las siguientes posibilidades de elección de β_{k+1},

- *β_{k+1} dado por (5) siendo*

$$s_k^T y_k = \max_i \nabla F_i(x_{k+1})^T s_k - \nabla F_i(x_k)^T s_k$$

  o

- β_{k+1} *dado por (5) siendo*
$$s_k^T y_k = \varphi_{x_{k+1}}(s_k) - \varphi_{x_k}(s_k).$$

Esta última elección, se reduce al coeficiente espectral (5) cuando $m = 1$.

Como sucede para el caso escalar, se puede probar que el Algoritmo termina en un número finito de iteraciones, con un punto estacionario, o genera una sucesión infinita de iterados no estacionarios.

Lema 2 *Sea* $\{x_k\} \subset \mathbb{R}^n$ *una sucesión generada por el algoritmo. Entonces,* $\{x_k\} \subset C$.

Teorema 2 *Cada punto de acumulación de una sucesión* $\{x_k\}$ *generada por el algoritmo es un punto factible estacionario.*

3. Conclusión y trabajo futuro

En este trabajo estudiamos alternativas para definir un método del tipo gradiente espectral proyectado para el problema multiobjetivo con restricciones. En este contexto proponemos dos posibles sucesiones de coeficiente espectral para definir la dirección de búsqueda. Se puede demostrar que el algoritmo propuesto hereda las propiedades de convergencia global que posee el método del gradiente espectral proyectado para el caso escalar [1] y el método del gradiente proyectado con búsqueda monótona para el problema multiobjetivo presentado en [3]. Sabemos que el uso del gradiente espectral proyectado en el caso escalar tiene un mejor desempeño cuando es asociado a búsqueda lineal no monótona. Es por este motivo que, como trabajo futuro, proponemos estudiar posibles búsquedas no monótonas que posean propiedades de convergencia global y analizar el desempeño de la nueva propuesta del punto de vista práctico.

Referencias

[1] E.G. Birgin, J.M. Martínez and M. Raydan. *Inexact spectral projected gradient methods on convex sets*, Vol. 23, Isse 4. pp.539-559, 2003.

[2] E.H. Fukuda, L.M. Graña Drummond and F. M. P. Raupp. *An external penalty-type method for multicriteria.* Aceptado en TOP, 2016.

[3] E.H. Fukuda, L.M. Graña Drummond. *A survey on multioblective descent methods*, 2014.

[4] L.M. Graña Drummond and A.N. Iusem. *A projected gradient method for vector optimization problems*, Computational Optimization and Applications, vol. 28, pp. 5-29, 2004.