Study on the Effectiveness of Classical Fritz John Conditions

Amrita Pal
Department of Mathematics,
IFTM University
Moradabad, India

Prashant Chauhan
Department of Mathematics,
IFTM University
Moradabad, India

Paras Bhatnagar
Department of IT
G.L. Bajaj ITM
Gr. Noida, India

ABSTRACT

The classical Fritz John conditions have been enhanced through the addition of an extra necessary condition, and their effectiveness has been significantly improved (for the case where X is a closed convex set, and Bertsekas and Ozdaglar [1] for the case where X is a closed set). In this paper we will use the following assumptions instead of smoothness and the assumption of existence of an optimal solution will retain.

Keywords

Fritz John conditions, lower semicontinuous functions, convex programming problem.

1. INTRODUCTION

Assumption: (Closedness) The functions f and $g_1,....,g_r$ are closed.

We note that f and $g_1,.....,g_r$ are closed if and only if they are lower semicontinuous on X, i.e., for each $\bar{x} \in X$, we have

$$f\left(\bar{x}\right) \le \liminf_{x\in X, x\to \bar{x}} f\left(x\right),$$

$$g_j\left(\bar{x}\right) \le \liminf_{x\in X,\, x\to \bar{x}} g\left(x\right),$$

$$j = 1,...,r,$$

Now we will prove the Fritz John conditions.

Lemma 1: Consider the convex problem (P) and assume

that $-\infty < q^*$. If μ^* is a dual optimal solution, then

$$\frac{q^* - f\left(x\right)}{\| g^+\left(x\right)\|} \le \| \mu^* \|, \quad \text{for all } x \in X \text{ that are infeasible.}$$

Proof: For any $x \in X$ that is infeasible, we have from the definition of the dual function that

$$q^* = q\left(\mu^*\right) \le f\left(x\right) + \mu^{*\prime} g\left(x\right) \le$$
$$f\left(x\right) + \mu^{*\prime} g^+\left(x\right) \le f\left(x\right) + \|\mu^*\| \, \|g^+\left(x\right)\|.$$

The preceding lemma shows that the minimum distance to the set of dual optimal solutions is an upper bound for the cost improvement/constraint violation ratio $(q^* - f(x))/\|g^+\left(x\right)\|$. The next proposition shows that, under certain assumptions including the absence of a duality gap, this upper bound is sharp, and is asymptotically attained by an appropriate sequence $\{x^k\} \subset X$.

Proposition 1: Let the convex problem (P) and x^* be an optimal solution Then there exists a FJ-multiplier $\left(\mu_0^*, \mu^*\right)$ satisfying the following condition (C_1).

(C_1) If $\mu^* \neq 0$, then there exists a sequence $\{x^k\} \subset X$ of infeasible points that converges to $x^*$and satisfies

$$f\left(x^k\right) \to f^*, g^+\left(x^k\right) \to 0, \qquad (1)$$

$$\frac{f^* - f\left(x^k\right)}{\left\| g^+\left(x^k\right)\right\|} \to \begin{cases} \| \mu^* \| / \mu_0^* & \text{if } \mu_0^* \neq 0, \\ \infty & \text{if } \mu_0^* = 0, \end{cases} \qquad (2)$$

$$\frac{g^+\left(x^k\right)}{\left\| g^+\left(x^k\right)\right\|} \to \frac{\mu^*}{\|\mu^*\|} \qquad (3)$$

Proof: For positive integers k and m , we consider the saddle function

$$L_{k,m}\left(x, \xi\right) = f\left(x\right) + \frac{1}{k^3}\left\|x - x^*\right\|^2 + \xi' g\left(x\right) - \frac{1}{2m}\| \xi \|^2$$

, $\xi \ge 0, L_{k,m}\left(x, \xi\right)$.

Furthermore, for a fixed $x, L_{k,m}\left(x, \xi\right)$ is negative definite quadratic in ξ. For each k , we consider the set

$$X^k = X \cap \left\{ x \mid \left\| x - x^* \right\| \le k \right\}.$$

Since f and g_j are closed and convex when restricted to X, they are closed, convex, and coercive when restricted to X^k. Thus, we can use the Saddle Point theorem to assert that $L_{k,m}$ has a saddle point over $x \in X^k$ and $\xi \ge 0.$. This saddle point is denoted by $\left(x^{k,m},\ \xi^{k,m}\right)$

The infimum of $L_{k,m}\left(x, \xi^{k,m}\right)$ over $x \in X^k$ is attained at $x^{k,m}$, implying that

$$\begin{aligned} & f\left(x^{k,m}\right) + \frac{1}{k^3}\left\| x^{k,m} - x^* \right\|^2 + \xi^{k,m'} g\left(x^{k,m}\right) \\ &= \inf_{x \in X^k} \left\{ f(x) + \frac{1}{k^3}\left\| x - x^* \right\|^2 + \xi^{k,m'} g(x) \right\} \\ &\le \inf_{x \in X^k,\ g(x) \le 0} \left\{ f(x) + \frac{1}{k^3}\left\| x - x^* \right\|^2 + \xi^{k,m'} g(x) \right\} \\ &\le \inf_{x \in X^k,\ g(x) \le 0} \left\{ f(x) + \frac{1}{k^3}\left\| x - x^* \right\|^2 \right\} \\ &= f\left(x^*\right) \end{aligned} \qquad (4)$$

Hence, we have

$$\begin{aligned} L_{k,m}\left(x^{k,m}, \xi^{k,m}\right) &= f\left(x^{k,m}\right) + \frac{1}{k^3}\left\| x^{k,m} - x^* \right\|^2 + \xi^{k,m'} g\left(x^{k,m}\right) - \frac{1}{2m}\left\| \xi^{k,m} \right\|^2 \\ &\le f\left(x^{k,m}\right) + \frac{1}{k^3}\left\| x^{k,m} - x^* \right\|^2 + \xi^{k,m'} g\left(x^{k,m}\right) \\ &\le f\left(x^*\right). \quad (5) \end{aligned}$$

Since $L_{k,m}\left(x^{k,m}, \xi\right)$ is quadratic in ξ, the supremum of $L_{k,m}\left(x^{k,m},\ \xi\right)$ over $\xi \ge 0$ is attained at

$$\xi^{k,m} = mg^+\left(x^{k,m}\right). \qquad (6)$$

This implies that

$$\begin{aligned} L_{k,m}\left(x^{k,m}, \xi^{k,m}\right) &= f\left(x^{k,m}\right) + \frac{1}{k^3}\left\| x^{k,m} - x^* \right\|^2 + \frac{m}{2}\left\| g^+\left(x^{k,m}\right) \right\|^2 \\ &\ge f\left(x^{k,m}\right) + \frac{1}{k^3}\left\| x^{k,m} - x^* \right\|^2 \\ &\ge f\left(x^{k,m}\right). \end{aligned} \qquad (7)$$

From Eqs. (5) and (7), we see that the sequence $\left\{x^{k,m}\right\}$, with k fixed, belongs to the set $\left\{ x \in X^k \mid f(x) \le f\left(x^*\right) \right\}$, which is compact, for each k, $L_{k,m}\left(x^{k,m},\ \xi^{k,m}\right)$ is bounded from above by $f\left(x^*\right)$, so

$$\limsup_{m \to \infty} g_j\left(x^{k,m}\right) \le 0, \qquad \forall\ j = 1, \ldots\ldots, r.$$

Therefore, by using the lower semicontinuity of g_j, we obtain $g\left(\bar{x}^k\right) \le 0,$ implying that $\bar{x}^k$ feasible solution of problem (P), so that $f\left(\bar{x}^k\right) \ge f\left(x^*\right).$ Using Eqs. (5) and (7) together with the lower semicontinuity of f, we also have

$$f\left(\bar{x}^k\right) \le \liminf_{m \to \infty} f\left(x^{k,m}\right) \le \limsup_{m \to \infty} f\left(x^{k,m}\right) \le f\left(x^*\right),$$

thereby showing that for each k,

$$\lim_{m \to \infty} f\left(x^{k,m}\right) = f\left(x^*\right)$$

Together with Eqs. (5) and (7), this also implies that for each k ,

$$\lim_{m \to \infty} x^{k,m} = x^*$$

Combining the preceding relations with Eqs.(5) and (7), for each k, we obtain

$$\lim_{m \to \infty} \left(f\left(x^{k,m}\right) - f\left(x^*\right) + \xi^{k,m'} g\left(x^{k,m}\right) \right) = 0 \qquad (8)$$

Denote

$$\delta^{k,m} = \sqrt{1 + \left\| \xi^{k,m} \right\|^2}, \quad \mu_0^{k,m} = \frac{1}{\delta^{k,m}}, \quad \mu^{k,m} = \frac{\xi^{k,m}}{\delta^{k,m}} \qquad (9)$$

dividing (8) by $\delta^{k,m}$, we obtain

$$\lim_{m \to \infty} \left(\mu_0^{k,m} f\left(x^{k,m}\right) - \mu_0^{k,m} f\left(x^*\right) + \mu^{k,m'} g\left(x^{k,m}\right) \right) = 0$$

By the preceding relations, for each k we can find a sufficiently large integer m_k such that

$$\left| \mu_0^{k,m_k} f\left(x^{k,m_k}\right) - \mu_0^{k,m_k} f\left(x^*\right) + \mu^{k,m_k'} g\left(x^{k,m_k}\right) \right| \le \frac{1}{k} \qquad (10)$$

and

$$\left\| x^{k,m_k} - x^* \right\| \le \frac{1}{k}, \left| f\left(x^{k,m_k}\right) - f\left(x^*\right) \right| \le \frac{1}{k}, \quad (11)$$
$$\left\| g^+\left(x^{k,m_k}\right) \right\| \le \frac{1}{k}$$

Dividing both sides of the first relation in Eq. (4) by δ^{k,m_k}, we obtain

$$\mu_0^{k,m_k} f\left(x^{k,m_k}\right) + \frac{1}{k^3 \delta^{k,m_k}} \left\| x^{k,m_k} - x^* \right\|^2$$
$$+ \mu^{k,m_k}{}' g\left(x^{k,m_k}\right)$$
$$\le \mu_0^{k,m_k} f(x) + \mu^{k,m_k}{}' g(x)$$
$$+ \frac{1}{k\,\delta^{k,m_k}} \qquad \forall x \in X^k,$$

also $\left\| x - x^* \right\| \le k$, $\forall x \in X^k$.

Without loss of generality, we will assume that the entire sequence $\left\{ \left(\mu_0^{k,m_k}, \mu^{k,m_k} \right) \right\}$ converges to $\left(\mu_0^*, \mu^* \right)$. Taking the limit as $k \to \infty$, and using Eq. (10), we obtain

$$\mu_0^* f\left(x^*\right) \le \mu_0^* f(x) + \mu^{*\prime} g(x), \forall x \in X.$$

Since $\mu^* \ge 0,$, this implies that

$$\begin{aligned} \mu_0^* f\left(x^*\right) &\le \inf_{x \in X} \left\{ \mu_0^* f(x) + \mu^{*\prime} g(x) \right\} \\ &\le \inf_{x \in X, g(x) \le 0} \left\{ \mu_0^* f(x) + \mu^{*\prime} g(x) \right\} \\ &\le \inf_{x \in X, g(x) \le 0} \mu_0^* f(x) \\ &= \mu_0^* f(x) \end{aligned}$$

Thus we have

$$\mu_0^* f\left(x^*\right) = \inf_{x \in X} \left\{ \mu_0^* f(x) + \mu^{*\prime} g(x) \right\},$$

$\left(\mu_0^*, \mu^* \right)$ satisfies (i).

If $\mu^* = 0,$, then $\mu_0^* \ne 0,$ (C_1) is automatically satisfied, and $\mu^* / \mu_0^* = 0$ has minimum norm.

Moreover, condition (i) yields $f^* = \inf_{x \in X} f(x),$ so that (C_1), is satisfied by only $\mu^* = 0$.

Assume now that $\mu^* \ne 0$, so that the index set $J = \left\{ j \ne 0 \mid \mu_j^* > 0 \right\}$ is nonempty. For large k, $\xi_j^{k,m_k} > 0$, $g_j\left(x^{k,m_k}\right) > 0$, $\forall j \in J$. Using Eqs. (6), (9) and the fact that $\mu^{k,m_k} \to \mu^*$ we obtain

$$\frac{g^+\left(x^{k,m_k}\right)}{\left\| g^+\left(x^{k,m_m}\right) \right\|} = \frac{\mu^{k,m_k}}{\left\| \mu^{k,m_k} \right\|} \to \frac{\mu^*}{\left\| \mu^* \right\|}$$

Using also Eq. (5) and $f\left(x^*\right) = f^*$, we have that

$$\frac{f^* - \left(x^{k.m_k}\right)}{\left\| g^+\left(x^{k.m_k}\right) \right\|} \ge \frac{\xi^{k,m_k\prime} g\left(\overleftarrow{x}^{k.m_k}\right)}{\left\| g^+\left(x^{k.m_k}\right) \right\|} = \left\| \xi^{k.m_k} \right\| = \frac{\left\| \mu^{k.m_k} \right\|}{\mu_0^{k,m_k}} \quad (12)$$

If $\mu_0^* = 0$, then $\mu_0^{k,m_k} \to 0$, so with $\left\| \mu^{k,m_k} \right\| \to \left\| \mu^* \right\| > 0$ we have

$$\frac{f^* - \left(x^{k.m_k}\right)}{\left\| g^+\left(x^{k.m_k}\right) \right\|} \to \infty$$

If $\mu_0^* \ne 0$, then together with $\mu_0^{k,m_k} \to \mu_0^*$ and $\left\| \mu^{k,m_k} \right\| \to \left\| \mu^* \right\|$ we have

$$\liminf_{k \to \infty} \frac{f^* - f\left(x^{k.m_k}\right)}{\left\| g^+\left(x^{k.m_k}\right) \right\|} \ge \frac{\left\| \mu^* \right\|}{\mu_0^*}$$

Using geometric multiplier μ^* / μ_0^* and $f^* = q^*$, Lemma 1 implies that μ^* / μ_0^* is of minimum norm. Hence, sequence $\left\{ x^{k,m_k} \right\}$ also satisfies conditions (1)-(3) of the proposition, concluding the proof.

2. Minimum-norm Dual Optimal solutions

Proposition 2: (Fritz John Conditions) Consider the convex problem (P), and assume that $f^* < \infty$. Then there exists a FJ-multiplier $\left(\mu_0^*, \mu^* \right)$.

Proof: If f = *-8*, then $\mu_0^* = 1$ and $\mu^* = 0$ form a FJ-multiplier. We may thus assume that f^* is finite. Consider the subset of R^{r+1} given by

$$M = \{ (u_1, ..., u_r, w) | \text{ there exists } x \in X \text{ such that } g_j(x) \le u_j,\ j = 1, ..., r,\ f(x) \le w \}$$

We first show that M is convex. Consider vectors $(u, w) \in M$ and $(\bar{u}, \bar{w}) \in M$, and we show that their convex combinations lie in M. The definition of M implies that for some $x \in X$ and $\tilde{x} \in X$, we have

$$f(x) \le w, \quad g_j(x) \le u_j, \quad j=1,....., r,$$
$$f(\tilde{x}) \le \tilde{w}, \quad g_j(\tilde{x}) \le \tilde{u}_j, \quad j=1,....., r,$$

For any $\alpha \in [0, 1]$, we multiply these relations with α and 1 - α, respectively, and add them. By using the convexity of f and g_j, we obtain

$$f(\alpha x + (1-\alpha)\tilde{x}) \le \alpha f(x) + (1-\alpha) f(\tilde{x}) \le \alpha w + (1-\alpha)\tilde{w},$$
$$g_j(\alpha x + (1-\alpha)\tilde{x}) \le \alpha g_j(x) + (1-\alpha) g_j(\tilde{x}) \le \alpha u_j + (1-\alpha) u_j,$$
$$j = 1, ..., r.$$

In view of the convexity of X, we have $\alpha x + (1-\alpha)\tilde{x} \in X$ *ax* +(1 - *a*)˜ *x X*, so these inequalities imply that the convex combination of (u, w) and $(\tilde{u}, \tilde{w})$, i.e., $(\alpha u + (1-\alpha)\tilde{u}, \alpha w + (1-\alpha)\tilde{w})$, belongs to M. This proves the convexity of M.

Therefore, there exists a hyper-plane passing through $(0, f^*)$ and containing M in one of its closed half spaces, i.e., there exists a vector $(\mu^*, \mu_0^*) \ne (0, 0)$ such that

$$\mu_0^* f^* \le u_0^* w + \mu^{*'} \mu, \qquad \forall (u, w) \in M. \quad (13)$$

This relation implies that

$$\mu_0^* \ge 0, \quad \mu_j^* \ge 0, \qquad \forall\, j = 1, ..., r,$$

since for each $(u, w) \in M$, we have that $(u, w+\gamma) \in M$ and

$$(u_1, ..., u_j + \gamma, ..., u_r, w) \in M \text{ for all } \gamma > 0 \text{ and j.}$$

Finally, since for all $x \in X$, we have g

$(g(x), f(x)) \in M$, Eq. (13) implies that

$$\mu_0^* f^* \le \mu_0^* f(x) + \mu^{*'} g(x), \qquad \forall\, x \in X.$$

Taking the infimum over all $x \in X$, it follows that

$$\begin{aligned} \mu_0^* f^* &\le \inf_{x \in X} \{ \mu_0^* f(x) + \mu^{*'} g(x) \} \\ &\le \inf_{x \in X, g(x) \le 0} \{ \mu_0^* f(x) + \mu^{*'} g(x) \} \\ &\le \inf_{x \in X, g(x) \le 0} \mu_0^* f(x) \\ &= \mu_0^* f^*. \end{aligned}$$

Hence above equality holds, that proves the result.

Lemma 2: Consider the convex problem (P), and assume that $f^* < \infty$.

For each $\delta > 0$, let

$$f^\delta = \inf_{\substack{x \in X \\ g_j(x) \le \delta,\ j=1,...,r}} f(x) \quad (14)$$

Then the dual optimal value q^* satisfies f = $f^\delta \le q^*$ for all $\delta > 0$ and

$$q^* = \lim_{\delta \downarrow 0} f^\delta.$$

Proof: We note that either lim $\delta \downarrow 0$ f^δ exists and is finite, or else lim $\delta \downarrow 0$ $f^\delta = -\infty$,

since f^δ is monotonically nondecreasing as $\delta \downarrow 0$, and $f^\delta \le f^*$ for all $\delta > 0$. Since $f^* < \infty$, there exists some $\bar{x} \in X$ such that $g(\bar{x}) \le 0$. Thus, for each $\delta > 0$ such that $f^\delta > -\infty$, the Slater condition is satisfied by Prop. 2 and the subsequent discussion, there exists a $\mu^\delta \ge 0$ satisfying.

$$f^{\delta} = \inf_{x \in X} \left\{ f(x) + \mu^{\delta'} g(x) - \delta \sum_{j=1}^{r} \mu_j^{\delta} \right\}$$

$$\le \inf_{x \in X} \left\{ f(x) + \mu^{\delta'} g(x) \right\}$$
$$= q\left(\mu^{\delta}\right)$$
$$\le q^*$$

For each $\delta > 0$ such that $f^{\delta} = -\infty$, we also have $f^{\delta} \le q^*$, so that

$$f^{\delta} \le q^*, \qquad \forall \delta > 0.$$

By taking the limit as $\delta \downarrow 0$, we obtain

$$\lim_{\delta \downarrow 0} f^{\delta} \le q^*$$

Consider (1) $f^{\delta} > -\infty$ for all $\delta > 0$ that are sufficiently small, and (2) $f^{\delta} > -\infty$ for all $\delta > 0$. In case (1), for each $\delta > 0$ with $f^{\delta} > -\infty$ choose $x^{\delta} \in X$ such that $g_j\left(x^{\delta}\right) \le \delta$ for all j and $f\left(x^{\delta}\right) + \delta$. Then, for any $\mu \ge 0$,

$$q(\mu) = \inf_{x \in X} \left\{ f(x) + \mu' g(x) \right\} \le$$
$$f\left(x^{\delta}\right) + \mu' g\left(x^{\delta}\right) \le$$
$$f^{\delta} + \delta + \delta \sum_{j=1}^{r} \mu_j$$

Taking the limit as $\delta \downarrow 0$, we obtain

$$q(\mu) \le \lim_{\delta \downarrow 0} f^{\delta}$$

so that $q^* \le \lim \delta \downarrow 0$, In case (2), choose $x^{\delta} \in X$ such that $g_j\left(x^{\delta}\right) \le \delta$ for all j and $\mathrm{f}\left(x^{\delta}\right) \le -1/\delta$ Then, similarly, for any $\mu \ge 0$, we have

$$q(\mu) \le f\left(x^{\delta}\right) + \mu' g\left(x^{\delta}\right) \le -\frac{1}{\delta} + \delta \sum_{j=1}^{r} \mu_j,$$

so by taking $\delta \downarrow 0$, we obtain $q(\mu) = -\infty$ for all $\mu \ge 0$, and hence also

$$q^* = -\infty = \lim \downarrow_0 f^{\delta}.$$

3. CONCLUSION

It is shown that the minimum distance to the set of dual optimal solutions is an upper bound for the cost improvement/constraint violation ratio ($q^* - f(x)) / \| g^+(x) \|$. Under the certain assumptions including the absence of a duality gap it is also shown that this upper bound is sharp, and is asymptotically attained by an appropriate sequence $\{x^k\} \subset X$.

4. REFERENCES

[1] Bertsekas, D. P., and Ozdaglar, A. E., 2002. “Pseudonormality and a Lagrange Multiplier Theory for Constrained Optimization,” J. Opt. Theory Appl., Vol. 114, 2002, pp. 287-343.

[2] John, F., 1948. “Extremum Problems with Inequalities as Subsidiary Conditions,” in Studies and Essays: Courant Anniversary Volume, K. O. Friedrichs, Neugebauer, O. E., and Stoker, J. J., (Eds.), Wiley-Interscience, N. Y., pp. 187-204.

[3] Rockafellar, R. T., 1970. Convex Analysis, Princeton Univ. Press, Princeton, N. J. 40.

[4] Song Xu, “A non-interior path following method for convex quadratic programming problems with bound constraints,” Comput. Optim. Applic., vol. 27, pp. 285–303, 2004.

[5] Treves, F. (1967), Locally Convex Spaces and Linear Partial Differential Equations, Springer Verlag, New-York.

[6] Truong, X.D.H. (1994), On the existence of efficient points in locally convex spaces, J.Global Optimization, 4, 265–278.

[7] Wang J.L., Zhang M.W., and Du T.S., “A primal-dual infeasible interior point algorithm for separable convex quadratic programming with box constraints,” J. Hebei Normal University, Natural Science Edition, vol. 26, no. 6, pp. 568–572, 587, 2002.