WORKING PAPER- TRANSPORTATION RESEARCH GROUP
Jan. 26.2014, UNIVERSITY OF MASSACHUSETTS LOWELL

THE NEAR BAYES NEAR MINIMAX SOLUTION FOR MULTI - OBJECTIVE OPTIMIZATION

Lee Jones & Nathan Gartner

I. SUMMARY

Given k cost functions F_i (x) corresponding to k outputs based on a control vector x, we try to simultaneously minimize the costs (or penalties). If this simultaneous minimization can be achieved with some control vector u, we say that "utopia" is attained by u. Usually this is not the case and sub-utopian solution vectors x* have been proposed in the literature [Marler and Arora 2004]. Two well known proposed control vectors are the minimax solution $x^* = x^m$ and the Pareto optimal $x^* = x^p$.

x^m minimizes the function $F^m (x) = \max_{i=1,2,\ldots k} \{ F_i (x) \}$ while x^p is a vector x* for which there is no other vector y such that $F_i(y) \leq F_i(x^*)$ for i = 1,2,…k and, for at least one j , $F_j(y) < F_j(x^*)$.

In general these two control vectors may differ: For instance consider $F_1 (x) = 1-x$ $x< 0$, 1 $0\leq x \leq 1$, and x $x > 1$. Let $F_2 (x) = 10^{-2} + 10^{-2} (x – 2)^2$. Then any x in [0,1) is minimax but not Pareto optimal while any x in (1,2] is Pareto optimal but not minimax. 1 is both minimax and Pareto optimal.

Another proposed control vector is the Bayes (or weighted) x^b, which minimizes $\sum_{i=1,..k} w_i F_i (x)$ for some probabilities w_i on {1,2,…k}. If each $w_i > 0$, then it is easily proved that x^b is Pareto optimal.

If however we assume that the control space is closed and convex and that each function is convex and has been rescaled by dividing by its minimum (a completely natural assumption for comparing relative costs (penalties)), the relationship between minimax and Pareto optimal and their properties becomes more interesting: From our theorem in Section II we will see that a minimax x* is always Pareto optimal for k=2 and, for arbitrary k, that a (at least) "binary equi-cost state " exists at a minimax x* in the non utopian case- namely that at least two of the F_i 's attain the minimax value at x* , for any neighborhood of x* two of those attaining the minmax value at x* take a larger value at some (different) point in the neighborhood while all of those attaining this minimax value at x* take smaller values at some (different) points in the neighborhood, and finally that one cannot guarantee the existence of three functions taking the minimax value. These assertions are , however, false if we weaken the assumption of convexity to unimodality. (But they are true for strictly unimodal functions , for which we give a different proof in Sec. III. Our proposed control vector, however, assumes convexity of the functions.)

Based on the appeal of the binary equi-cost condition we propose to apply minimaxity to two functions derived from the original k (since we can always identify the two "participants " in the equi-cost state). In light of the previous discussion one of them might be $G_2 (x) = F^m (x) / m$ where m is the minimum of $F^m (x)$. (Note that F^m is convex since each F_i is convex.) As weighted costs are often considered as important by users the other is $G_1 (x) = (1/b) \sum_{i=1,..k} w_i F_i (x)$ where w_i are probabilities assigned to the various control functions and b is the minimum value of the weighted sum $\sum_{i=1,..k} w_i F_i (x)$. Hence the near Bayes near minimax solution x^{nbnm} is that which minimizes $\max\{G_1(x),G_2(x)\}$. In non-utopian cases this has the equilibrium properties described above and, if all $w_i > 0$, the solution x^{nbnm} is Pareto optimal but not minimax for the original k function problem. We note that approaches optimizing weighted sums of the form $\gamma G_2 (x) + \rho G_1 (x)$ have been proposed in the

literature where the parameters γ,ρ are user specified. The NBNM solution can be reduced to this form but γ,ρ are unpredictable and solving nbnm directly, as we pose it, is often an unconstrained convex programming problem.

II. MATHEMATICAL PROOFS OF ASSERTIONS

Theorem: For I =1,2,…k, let F_i (x) be convex positive real valued functions on a closed convex subset of d-dimensional Euclidean space with each F_i having a minimum value of 1. Assume no x simultaneously minimizes all control functions (utopia is not attained). Then at any minimax x* (which minimizes $F^m(x)$)

a. At least two F_i 's attain the minimax value $F^m(x^*)$ at x*. Consider all F_i for which the minimax value is attained at x*.

b. Given any convex neighborhood N of x*: All functions in a. take on smaller values than $F^m(x^*)$ at some (possibly different) points in the neighborhood N. At least two of the functions in a. take a larger value than $F^m(x^*)$ at some (possibly different) point in the neighborhood N.

<u>proof</u>: Suppose only one function, say F_1, attains the minimax value at x*. Since F_1 attains its minimum value 1 at say x_1, $F_1 (\alpha x_1 + (1-\alpha)x^*) < F^m(x^*)$ for $0 < \alpha < 1$. For sufficiently small α, $F_l (\alpha x_1 + (1-\alpha)x^*) < F^m(x^*)$ for l = 2,3,… k. Hence $F^m(x^*)$ is not the minimax value, a clear contradiction. So a. holds. (Example 1. below shows that we cannot always find three or more such functions.)

Repeating the above argument for any F_i in a. (with $F_i(x_i) = 1$), we see that $F_i (\alpha_i x_i + (1-\alpha_i)x^*) < F^m(x^*)$ for a sufficiently small α_i such that $\alpha_i x_i + (1-\alpha_i)x^*$ lies in N. Now suppose $F_i (x) \leq F^m(x^*)$ for all F_i attaining $F^m(x^*)$ at x* and all x in N. Let x** be the sample mean of the points $\alpha_i x_i + (1-\alpha_i)x^*$. By the convexity of N, x** lies in N and each $F_i (x^{**}) < F^m(x^*)$. By taking the α_i sufficiently small we can ensure that x** is sufficiently close to x* that $F_j (x^{**}) < F^m(x^*)$ for those F_j not attaining the minimax value in a. But then $F^m(x^*)$ is not the minimax value, a clear contradiction. Hence at least one such F_i takes a bigger value than $F^m (x^*)$ at some point in N. Suppose now that there is only one such F_i, say F_l. Now repeat the latter construction except replace the sample mean by a proper convex combination x** of the points $\alpha_i x_i + (1-\alpha_i)x^*$ that puts sufficient weight on the point $\alpha_l x_l + (1-\alpha_l)x^*$ to ensure that $F_l (x^{**}) < F^m(x^*)$. As before $F^m(x^*)$ is not the minimax value, a clear contradiction. So we must have at least two such functions taking larger values in N. **Q.E.D.**

(Example 2. below demonstrates that the theorem is false if we weaken the convexity assumption to that of unimodal functions.)

Corollary I: For k = 2, any minimax x* is Pareto optimal assuming each F_i is convex with minimum 1.

<u>proof</u>: If utopia is attained then x* is clearly Pareto optimal since both functions equal 1 at x*. Otherwise both take the minimax value at x* by the theorem. Now if x* is not Pareto optimal then there

exists a x^{**} with say $F_1(x^{**}) < F^m(x^*)$ and $F_2(x^{**}) \leq F^m(x^*)$. Since $F^m(x^*)$ is the minimax value we must have $F_2(x^{**}) = F^m(x^*)$. Hence x^{**} is also minimax but a. of the theorem fails to hold. By contradiction x^* must be Pareto optimal. **Q.E.D.**

(Example 3. below shows that Corollary I is false for k = 3.)

Corollary II: Assuming each F_i is convex with minimum 1 and all $w_i > 0$, x^{nbnm} is Pareto optimal for the k function problem.

proof: if not there is a control vector y for which $F_i(y) \leq F_i(x^{nbnm})$ for i = 1,2,…k and, for at least one j , $F_j(y) < F_j(x^{nbnm})$. Since $w_j > 0$, $G_1(y) < G_1(x^{nbnm})$. Clearly $G_2(y) \leq G_2(x^{nbnm})$ and, by the min max property of x^{nbnm} , $G_2(y) = G_2(x^{nbnm})$. Hence y solves the NBNM problem but a. of the theorem fails to hold, a clear contradiction. **Q.E.D.**

Example 1 : Take $F_1(x) = 1$, $F_2(x) = 1 + (x - 1)^2$, and $F_3(x) = 1 + (x + 1)^2$. The minmax value is 2 , the minimax solution x^* is uniquely 0 but only F_2 and F_3 take the value 2 at 0.

Example 2: Take $F_1(x) = 1 + x^2$ and $F_2(x) = 2 – x$ for $x \leq 0$, 2 for $0 < x < 1$, $1 + (x – 2)^2$ for $x \geq 1$. Then x = .5 is minimax but only F_2 attains the minimax value at .5.

Example 3: Take $F_1(x,y) = 1 + 10^{-2} (y - .1)^2$, $F_2(x,y) = 1 + (x - 1)^2$, and $F_3(x,y) = 1 + (x + 1)^2$. The minmax value is 2 , a minimax solution is $(x^*, y^*) = (0,0)$ but , at (0, .1), F_1 takes a smaller value than at (0,0) while F_2 and F_3 take the same value as at (0,0). Hence (0,0) is not Pareto optimal.

III. EXTENSION OF THEOREM TO STRICTLY UNIMODAL FUNCTIONS:

We use the following definitions for unimodality: A continuous function on a closed convex set is unimodal if there is a point in the domain such that on the intersection of the domain and any line through that point the function decreases (not necessarily strictly) and then increases (not necessarily strictly) while achieving its minimum at the point. The function is strictly unimodal if the latter increases and decreases are both strict.

The theorem is true for strictly unimodal functions. Namely, in any neighborhood of a non-utopian x^*, every function achieving the minimax value takes a smaller value in the neighborhood.(Consider the behavior of the function on the line through the point minimizing the function and x^* .) Hence the first assertion in b. holds. Now clearly, if only one function achieved the minimax value at x^* then we could find a smaller minimax value for some point x^{**} in the neighborhood, which is impossible. So a. is valid.

The second assertion in b. , finding two points achieving the minimax value with larger values assumed in the neighborhood, requires a slightly different approach as in the theorem:

Let us first assume, for some neighborhood N of x^*, all functions F_1 , F_2 , ….. F_s (there are at least two such , so $s \geq 2$) taking the minimax value at x^* take values less than or equal the minimax value in N. (We can also assume, by making N sufficiently smaller if necessary, that all other functions Fi take smaller values than the minimax value everywhere in N.) Somewhere in N, F_1 takes a smaller value; in fact it takes smaller values everywhere in a proper sub neighborhood N' of N but these values may be strictly bounded below by 1. Now F_2 takes a smaller value somewhere in N' and everywhere in a proper sub neighborhood N''of N' and these values of F_2 can be strictly bounded below in N'' by 1. Continuing the argument we eventually find a subneighborhood N* of N where all s functions take values less than the minimax value while all other F_i's also take smaller values in N*. This is a clear contradiction.

Finally, we assume there is only one function such F_l above taking bigger values in the N constructed. We can still find a point in N where F_l takes a smaller value and repeat the construction exactly as above arriving at the same contradiction. **QED.**