JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 113, No. 2, pp. 399–407, May 2002 (©2002)

TECHNICAL NOTE

Universal Alignment Probability Revisited[1]

S. Y. LIN[2] AND Y. C. HO[3]

Abstract. In this note, we quantify and validate the representativeness of the uniformly sampled set N for the search space Θ and the use of universal alignment probability (UAP) curves.

Key Words. Ordinal optimization, alignment probability, good enough subset, selected subset, ordered performance curve.

1. Introduction

One of the major claims of the ordinal optimization (OO) approach is this: By using a crude model to quickly and roughly estimate the performance of a set of choices, one can under some fairly general conditions determine good enough choices with high probability. This probability, defined as alignment probability, gives the chance that the estimated good enough choices are indeed truly good enough choices. The main application of OO is to quantify various quick-and-dirty or back of the envelope schemes for computationally complex design problems. Mathematically, we use the following definitions:

Θ = search space for optimization;
θ = elements of Θ;
N = set of uniformly samples elements of Θ, usually of size sufficiently large, say 1000;
GS_N = good enough subset of N, typically the top n%;

[1]The work in this paper was partially supported by Army Contracts DAAL03-92-G-0115 and DAAH04-95-0148, by Air Force Grant F49620-98-1-0387, by Electric Power Research Institute Contract WO8333-03, and by National Science Council in Taiwan, ROC.

[2]Professor, Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, ROC.
[3]Research Professor, Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts.

GS_Θ = good enough subset of Θ, unless otherwise specified, we use the same top $n\%$ as definition;

SS = selected subset of N (estimated top $n\%$ by using a crude model).

Then, we define

$$\text{alignment probability} \equiv P_A \equiv \text{Prob}\{|\text{SS} \cap \text{GS}_N| \geq \text{k}\}. \quad (1)$$

The P_A are usually referred to as universal alignment probabilities (UAP), since they are applicable to all optimization problems within five broad categories (Ref. 1) and have been used successfully in many practical complex optimization problems (Refs. 2–5).

One implicit assumption in OO theory is that, for sufficiently large N, the set N is more or less representative of the search space Θ, i.e., GS_N is representative of GS_Θ. But representativeness does not necessarily imply, for example, $GS_N \subseteq GS_\Theta$, even though the probability that this is true can be high. In this note, we answer the following question: what is the relation between the definition (1) and the definition (2) below,

$$\text{Alignment Probability*} \equiv P_A^* \equiv \text{Prob}\{|\text{SS} \cap \text{GS}_\Theta| \geq k^*\}. \quad (2)$$

Intuitively,

$$P_A^* < P_A, \qquad \text{for } k^* = k,$$

since $GS_N \subseteq GS_\Theta$ is not always true. Alternatively, we can define

GS_N^r = reduced good enough subset of N, i.e., the top $m\%$, where $m < n$.

The idea is that, by requiring a more stringent definition of what is good enough in N, we guarantee with high probability that $GS_N^r \subseteq GS_\Theta$. Consequently, using the tabulated value of P_A of (1) in Ref. 1, we can in fact lower bound P_A^* in (2) with little extra work. The purpose of this note is to establish this bound and quantify what heretofore has been only empirical evidence of successful uses of the OO approach.

2. Analysis and Results

In this section, we will derive a lower bound of P_A^* and determine the size of *SS* for the desired alignments k^* with a high alignment probability P_A^*.

2.1. Lower Bound of P_A^*. First, we define the notations for the following events:

$$E_1 \equiv \{|\text{SS} \cap \text{GS}_N| \geq k\},$$

$$E_2 \equiv \{\text{GS}_N^r \subseteq \text{GS}_\Theta\},$$

$$E_3 \equiv \{|\{k\} \cap \mathrm{GS}_N^r| \geq k^*\},$$

$$E_4 = \{|\mathrm{SS} \cap \mathrm{GS}_N^r| \geq k^*\},$$

$$E_5 \equiv \{|\mathrm{SS} \cap \mathrm{GS}_\Theta| \geq k^*\},$$

where the set $\{k\}$ in E_3 is composed of the top k designs of $\mathrm{SS} \cap \mathrm{GS}_N$ in E_1. Clearly,

$$P_A = \mathrm{Prob}\{E_1\}, \qquad P_A^* = \mathrm{Prob}\{E_5\},$$

and the following inequality holds:

$$P_A^* \geq \mathrm{Prob}\{E_1 \text{ and } E_2 \text{ and } E_5\}. \tag{3}$$

By the definition of conditional probability, we have

$$\mathrm{Prob}\{E_1 \text{ and } E_2 \text{ and } E_5\} = \mathrm{Prob}\{E_1 \text{ and } E_2\} \times \mathrm{Prob}\{E_5 | E_1 \text{ and } E_2\}. \tag{4}$$

Since the alignments between SS and GS_N have nothing to do with the containment of GS_N^r in GS_Θ, the events E_1 and E_2 are independent of each other. Thus, (4) becomes

$$\mathrm{Prob}\{E_1 \text{ and } E_2 \text{ and } E_5\}$$

$$= \mathrm{Prob}\{E_1\} \times \mathrm{Prob}\{E_2\} \times \mathrm{Prob}\{E_5 | E_1 \text{ and } E_2\}. \tag{5}$$

Suppose that the events E_1 and E_2 are given; then, event E_3 implies E_4, because

$$\{k\} \cap \mathrm{GS}_N^r \subseteq (\mathrm{SS} \cap \mathrm{GS}_N) \cap \mathrm{GS}_N^r = \mathrm{SS} \cap \mathrm{GS}_N^r,$$

and event E_4 implies E_5, because

$$\mathrm{SS} \cap \mathrm{GS}_N^r \subseteq \mathrm{SS} \cap \mathrm{GS}_\Theta.$$

Hence, we have

$$\mathrm{Prob}\{E_4 | E_1 \text{ and } E_2\} \geq \mathrm{Prob}\{E_3 | E_1 \text{ and } \mathrm{E}_2\}, \tag{6}$$

$$\mathrm{Prob}\{E_5 | E_1 \text{ and } E_2\} \geq \mathrm{Prob}\{E_4 | E_1 \text{ and } E_2\}. \tag{7}$$

From (5)–(7), we can obtain

$$\mathrm{Prob}\{E_1 \text{ and } E_2 \text{ and } E_5\}$$

$$\geq \mathrm{Prob}\{E_1\} \times \mathrm{Prob}\{E_2\} \times \mathrm{Prob}\{E_3 | E_1 \text{ and } E_2\}. \tag{8}$$

Since E_3 is independent of E_2, we have

$$\mathrm{Prob}\{E_3 | E_1 \text{ and } E_2\} = \mathrm{Prob}\{E_3 | E_1\}. \tag{9}$$

Then, from (3), (8), and (9), we can obtain a lower bound of P_A^* as follows:

$$P_A^* \geq \mathrm{Prob}\{E_1\} \times \mathrm{Prob}\{E_2\} \times \mathrm{Prob}\{E_3 | E_1\}. \tag{10}$$

By the definition of E_1, E_2, and E_3, the result (10) says that, if the following three events occur:

$$\{|\mathrm{SS}\cap \mathrm{GS}_N|\geq k\}, \{\mathrm{GS}_N^r\subseteq \mathrm{GS}_\Theta\}, \quad \text{and} \quad \{|\{k\}\cap \mathrm{GS}_N^r|\geq k^*||\mathrm{SS}\cap \mathrm{GS}_N|\geq k\},$$

then the event $\{|\mathrm{SS}\cap \mathrm{GS}_\Theta|\geq k^*\}$ must occur. This is of course intuitively true. Now, since the tabulated value of $\mathrm{Prob}\{E_1\} = P_A$ has been given already in Ref. 1, to calculate the lower bound for P_A^*, we still need to calculate $\mathrm{Prob}\{\mathrm{GS}_N^r\subseteq \mathrm{GS}_\Theta\}$ and $\mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r|\geq k^*||\mathrm{SS}\cap \mathrm{GS}_N|\geq k\})$ in the following subsections.

2.2. Calculation of Prob{$\mathrm{GS}_N^r\subseteq \mathrm{GS}_\Theta$}. Denoting Prob($i$) as the probability that i designs in N are contained in GS_Θ, then

$$\mathrm{Prob}\{\mathrm{GS}_N^r\subseteq \mathrm{GS}_\Theta\} = 1 - \sum_{i=0}^{|GS_N^r|-1} \mathrm{Prob}(i). \tag{11}$$

The formula for Prob(i) under the assumption of uniform selection for N can be derived as follows.

Since there are $C(|N|, i)$ possible combinations for i designs of N contained in GS_Θ, and since each combination has the same probability, thus Prob(i) can be calculated by $C(|N|, i)$ times the probability that the first i designs of N are contained in GS_Θ. Hence, we have

$$\begin{aligned}\mathrm{Prob}(i) = C(|N|, i) \times \frac{|\mathrm{GS}_\Theta|}{|\Theta|} \times \frac{|\mathrm{GS}_\Theta|-1}{|\Theta|-1} \times \cdots \times \frac{|\mathrm{GS}_\Theta|-(i-1)}{|\Theta|-(i-1)} \\ \times \frac{|\overline{\mathrm{GS}_\Theta}|}{|\Theta|-i} \times \frac{|\overline{\mathrm{GS}_\Theta}|-1}{|\Theta|-(i+1)} \times \cdots \times \frac{|\overline{\mathrm{GS}_\Theta}|-(|N|-i)}{|\Theta|-|N|},\end{aligned} \tag{12}$$

where $\overline{\mathrm{GS}_\Theta}$ denote the complement of GS_Θ in the set Θ. Since $|\mathrm{GS}_\Theta|$, $|\overline{\mathrm{GS}_\Theta}|$, and $|\Theta|$ are very large, the subtracted terms in (12) are negligible. Consequently, (12) can be approximated as

$$\mathrm{Prob}(i) \cong C(|N|, i) \times \left(\frac{|\mathrm{GS}_\Theta|}{|\Theta|}\right)^i \times \left(\frac{|\overline{\mathrm{GS}_\Theta}|}{|\Theta|}\right)^{|N|-i}. \tag{13}$$

Remark 2.1. Consider a standard optimization problem,

$$\min_{\theta\in\Theta} J(\theta),$$

where Θ is the design space and $J(\cdot)$ is a performance measure defined on the design space. The ordinal performance curve (OPC) of a collection of ordered designs $\{\theta_1, \theta_2, \ldots, \theta_I\}$ selected from Θ defined in Ref. 1 is determined by the spread of the order performance $J_{[1]}, J_{[2]}, \ldots, J_{[I]}$, $J_{[i]}$ denotes

$J(\theta_i)$. Without loss of generality, the $J_{[i]}$ can be normalized into the range [0, 1]; i.e., for $i = 1, \ldots, 1$,

$$y_i = (J_{[i]} - J_{[1]})/(J_{[I]} - J_{[1]}).$$

Meanwhile, the ordered designs, spaced equally, are also mapped into the range [0, 1] such that, for all $i = 1, \ldots, I$,

$$x(\theta_{[i]}) \equiv x_i = (i-1)/(I-1).$$

There are five broad categories of OPC models:

(i) lots of good designs;
(ii) lots of intermediate but few good and bad designs;
(iii) equally distributed good, bad, and intermediate designs;
(iv) lots of good and lots of bad designs but few intermediate ones;
(v) lots of bad designs.

A graphical expression for these five OPC models is shown in Fig. 1.

2.3. Calculation of $\mathrm{Prob}\{|\{k\} \cap GS_N^r| \geq k^* | \, |SS \cap GS_N| \geq k\}$. Under the condition

$$|SS \cap GS_N| \geq k,$$

the probability for $|\{k\} \cap GS_N^r| \geq k^*$ will be higher if $k^* \leq k$ is smaller. Thus, it would be interesting to know what is the least value of k such that

$$\mathrm{Prob}\{|k\} \cap GS_N^r \geq |k^*| \, |SS \cap GS_N| \geq k\}$$

can be equal or close to 1. The advantage for having a smaller k is that we can achieve

$$|SS \cap GS_\Theta| \geq k^*$$

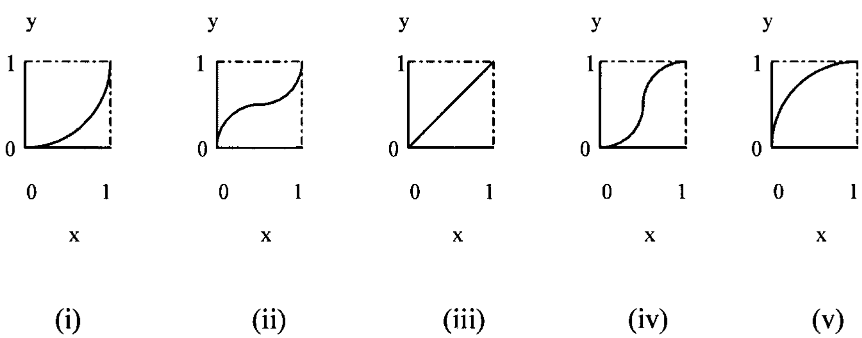

Fig. 1. The five OPC models.

with a high P_A^* by a smaller |SS| as can be observed from the term $\text{Prob}\{E_1\} = P_A$ in the lower bound of (10). Smaller |SS| for achieving the same P_A^* is of course favorable in the aspect of reducing computational complexity using the OO approach.

Unlike $\text{Prob}\{GS_N^r \subseteq GS_\Theta\}$, it is not suitable to use the assumption of uniform selection here, because $\{k\}$ is composed of the top k designs of $SS \cap GS_N$. Therefore, the performance values of the designs in $\{k\}$ are crucial in determining their order within GS_N and their intersection with GS_N^r.

Since both $\{k\}$ and GS_N^r are subsets of GS_N, without loss of generality we can employ a neutral-type OPC (Remark 2.1 and Ref. 1) to represent the normalized ordered performance of the equally spaced normalized ordered designs of GS_N and serve as the reference for placing the order of the designs of GS_N^r and $\{k\}$ within GS_N. Clearly, GS_N^r is composed of the top $(m/n) \times 100\%$ designs in GS_N. As shown in Fig. 2, the 45° straight-line represents the OPC of GS_N, and any design of GS_N with normalized performance less than m/n belongs to GS_N^r. To determine the order of the designs of $\{k\}$ within the set GS_N, we have to consider wider possibilities of the performances of the designs in $\{k\}$ within GS_N. Since the designs in $\{k\}$ are the top k designs of $SS \cap GS_N$, they are ranked higher in GS_N; therefore, their performances within GS_N should be categorized to the OPC classes of

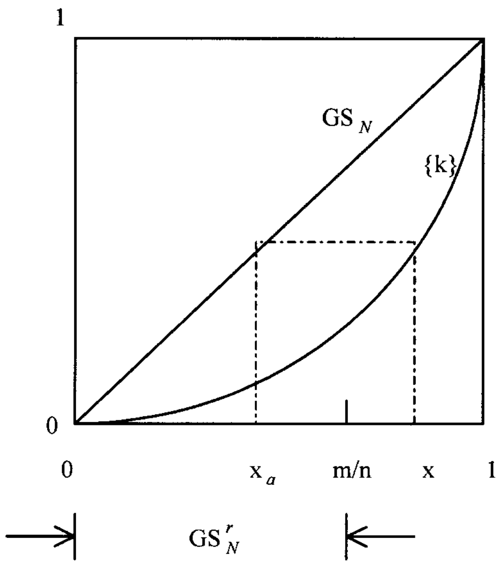

Fig. 2. Order adjustment for a sample x in $\{k\}$.

designs (i), (ii), and at least (iii). To cover a wider range of OPCs, we choose 37 different standardized OPCs belonging to the classes (i), (ii), (iii) formed from the combinations of the following parameters:[4]

$$\alpha = \{0.15, 0.25, 0.4, 0.55, 0.7, 0.85, 1.0, 1.5, 2.0, 3.0, 4.0\},$$

$$\beta = \{1.0, 2.0, 3.0, 4.0\},$$

as the domain of the OPCs of the set $\{k\}$. Now, for the designs of the set $\{k\}$ with a given OPC, we will determine their order within GS_N by two steps.

We first uniformly place them in the normalized ordered interval [0, 1] by the following manner: dividing [0, 1] into k subintervals and placing one design in the middle of a subinterval for all k designs. Then, we will adjust the order of these k designs based on their performances within the set GS_N. For example, if the given OPC of $\{k\}$ is a rather flat one, which corresponds to the class (i) of lots of good designs, as shown in Fig. 2, and if x, in the same figure, represents one of the uniformly placed k designs in [0, 1], then the order of the design x will be adjusted to that of x_a in GS_N, which has the same performance as x in $\{k\}$ as shown in Fig. 2.

Since all we care here is the intersection of $\{k\}$ and GS_N^r, thus if the performance of design in $\{k\}$ for a given OPC is less than m/n, this design belongs to GS_N^r. This means we do not need to execute the order adjustment step for the design x in $\{k\}$ described above. Based on this procedure for determining the intersection between $\{k\}$ and GS_N^r, we carry out 50,000 realizations of the OPCs, which are randomly selected from the 37 OPCs assigned to $\{k\}$ and obtain the relationship between k, k^*, and

$$\mathrm{Prob}\{|\{k\} \cap \mathrm{GS}_N^r| \geq \mathrm{k}^* \,|\, |\mathrm{SS} \cap \mathrm{GS}_N| \geq k\}$$

for $m/n = 0.7$, $n = 5$, $|N| = 1000$ shown in Tables 1 and 2. Comparing Table 1 with Table 2, we see that Table 1 provides us the least value of k for a given k^* such that

$$\mathrm{Prob}\{|\{k\} \cap \mathrm{GS}_N^r| \geq k^* \,\|\, \mathrm{SS} \cap \mathrm{GS}_N| \geq k\} = 1,$$

[4]To accommodate the five normalized OPC types described in Remark 2.1 using the smallest number of parameters, the inverse mapping of the incomplete beta function, parametrized by a pair of numbers α and β, is employed (Ref. 1). More precisely, the standardized OPC is determined by a two-parameter smooth curve,

$$\Lambda(x\,|\,\alpha, \beta) = F^{-1}(x\,|\,\alpha, \beta) = F(x\,|\,1/\alpha, 1/\beta),$$

where $F(x|\cdot, \cdot)$ is the incomplete beta function of the two parameters $(\cdot, \cdot)$. In general, $\alpha < 1, \beta > 1$ correspond to the OPC of type (i); $\alpha > 1, \beta > 1$ correspond to the OPC of type (ii); $\alpha = 1, \beta = 1$ correspond to the OPC of type (iii); $\alpha < 1, \beta < 1$ correspond to the OPC of type (iv); $\alpha > 1, \beta < 1$ correspond to the OPC of type (v).

Table 1. Relationship between k and k^* for $\mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS}\cup \mathrm{GS}_N| \geq k\} = 1$.

k	1	2	3	4	5	6	7	8	9	10
k^*	1	1	2	3	4	4	5	6	7	7
P_{kGr}	1	1	1	1	1	1	1	1	1	1

Notation: $P_{kGr} \equiv \mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS}\cap \mathrm{GS}_N| \geq k\}$.

Table 2. Relationship between k and k^* for $\mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS}\cap \mathrm{GS}_N| \geq k\} \leq 1$.

k	1	2	3	4	5	6	7	8	9	10
k^*	1	2	3	4	5	5	6	7	8	8
P_{kGr}	1	0.93	0.86	0.82	0.80	0.96	0.86	0.89	0.87	0.96

Notation: $P_{kGr} \equiv \mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS}\cap \mathrm{GS}_N| \geq k\}$.

and this is the answer to the question we posed at the beginning of this subsection.

2.4. Numerical Value for the Lower Bound of P_A^*. From (11) and (13), it can be computed easily that

$$\mathrm{Prob}\{\mathrm{GS}_N^r \subseteq \mathrm{GS}_\Theta\} = 0.991$$

if

$$m/n = 0.70, \qquad n = 5, \qquad |N| = 1000, \qquad |\Theta| = 10^8.$$

Thus, given that $P_A = 0.95$ and $\mathrm{Prob}\{\mathrm{GS}_N^r \subseteq \mathrm{GS}_\Theta\} = 0.991$, we have from (10) that

$$P_A^* \geq 0.942$$

if the values of k and k^* are chosen having

$$\mathrm{Prob}\{|\{k\}\cap \mathrm{GS}_N^r| \geq \mathrm{k}^* \| \mathrm{SS}\cap \mathrm{GS}_N| \geq \mathrm{k}\} = 1$$

as indicated in Table 1.

2.5. Determining the Size of SS for the Desired Alignment k^* with a High Alignment Probability P_A^*. Setting

$$m/n = 0.70, \qquad n = 5, \qquad |N| = 1000, \qquad |\Theta| = 10^8,$$

from Section 2.4, we have that

$$P_A^* \geq 0.942, \qquad \text{if } P_A = 0.95 \text{ and } \mathrm{Prob}\{|\{k\} \cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS} \cap \mathrm{GS}_N| \geq \mathrm{k}\} = 1.$$

Thus, to determine the size of SS for a desired k^* with $P_A^* \geq 0.942$, we first look up the corresponding smallest value of k with

$$\mathrm{Prob}\{|\{k\} \cap \mathrm{GS}_N^r| \geq k^* \| \mathrm{SS} \cap \mathrm{GS}_N| \geq k\} = 1$$

from Table 1. Then, for this k and $P_A = 0.95$, we can compute the size of SS by the function $Z(k, g)$, where g is the size of GS_N, given in (Ref. 1). Similar simple calculations can be done for m/n equal to a ratio other than 0.7. The bottom line is: If we choose a more restricted definition of GS_N about 30% less than GS_Θ, we can use the UAP curves as if they were designed to predict

$$P_A^* \equiv \mathrm{Prob}\{|\mathrm{SS} \cap \mathrm{GS}_\Theta| \geq \mathrm{k}\}$$

in (2), rather than P_A in (1).

3. Conclusions

This note quantifies and validates the notion that N is representative of Θ and the use of the universal alignment curves (UAP curves) first reported in Ref. 1.

References

1. Lau, T. W. E., and Ho, Y. C., *Universal Alignment Probabilities and Subset Selection for Ordinal Optimization*, Journal of Optimization Theory and Applications, Vol. 39, pp. 455–489, 1997.
2. Wieseltheir, J. E., Barnhart, C. M., and Ephremides, A., *Ordinal Optimization of Admission Control in Wireless Multihop Voice/Data Network via Standard Clock Simulation*, Journal of Discrete Event Dynamic Systems, Vol. 5, pp. 243–280, 1995.
3. Yang, M. S., Lee, L. H., and Ho, Y. C., *On Stochastic Optimization and Its Applications to Manufacturing*, Lectures in Applied Mathematics, Applied Mathematics Society-SIAM, Vol. 33, pp. 317–331, 1997.
4. Cassandras, C. G., Dai, L. Y., and Panayiotou, C. G., *Ordinal Optimization for a Class of Deterministic and Stochastic Discrete Resource Allocation Problems*, IEEE Transactions on Automatic Control, Vol. 43, pp. 881–900, 1998.
5. Ho, Y. C., Cassandras, C. C., Chen, C. H., and Dai, L. Y., *Ordinal Optimization and Simulation*, Journal of the Operations Research Society, Vol. 21, pp. 490–500, 2000.