Monitoring an Information Source under a Politeness Constraint – Online Supplement

Jonathan Eckstein
Department of Management Science and Information Systems and RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854 USA, jeckstei@rutcor.rutgers.edu
Avigdor Gal, Sarit Reiner
Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel,{avigal@ie,reiner@techunix}.technion.ac.il

A. Online Supplement

A.1. Proof of Proposition 1

Proof.

$$\begin{aligned} \mathrm{E}\big[U(s,f)\big] &= \mathrm{E}_{Q(s,f),R(s,f)}\big[U(s,f)\big] = \mathrm{E}_{Q(s,f)}\Big[\mathrm{E}_{R(s,f)}\big[U(s,f)\big]\Big] \\ &= \mathrm{E}_{Q(s,f)}\left[\sum_{q\in Q(s,f)} \mathrm{E}_{R(s,f)}\big[u(q,s)\big]\right] = \mathrm{E}_{Q(s,f)}\left[\sum_{q\in Q(s,f)} \Lambda(s,q)\right]. \end{aligned} \tag{1}$$

Under the assumptions on the query process, the number of queries $|Q(s,f)|$ during $(s,f]$ is a Poisson random variable with expected value $A(s,f) = \int_s^f a(\tau)d\tau$. Furthermore, if the number of queries $|Q(s,f)|$ is given, the conditional times $q \in Q(s,f)$ of the individual queries are independently distributed with probability density $d_q(\tau) = a(\tau)/A(s,f)$ over $(s,f]$. Combining these observations with (1), we obtain

$$\begin{aligned} \mathrm{E}\big[U(s,f)\big] &= \mathrm{E}_{Q(s,f)}\left[\sum_{q\in Q(s,f)} \Lambda(s,q)\right] = \sum_{k=0}^{\infty} \mathrm{P}\{|Q(s,f)| = k\}\Big(k\cdot \mathrm{E}_q\big[\Lambda(s,q)\big]\Big) \\ &= \left(\sum_{k=0}^{\infty} \mathrm{P}\{|Q(s,f)| = k\}\cdot k\right)\left(\int_s^f d_q(\tau)\Lambda(s,\tau)\,d\tau\right) \\ &= A(s,f)\int_s^f \frac{a(\tau)}{A(s,f)}\Lambda(s,\tau)\,d\tau = \int_s^f a(\tau)\left(\int_s^\tau \lambda(t)\,dt\right)d\tau, \end{aligned}$$

which closely resembles (2). Indeed, defining the set $Z(s,f) = \{(t,\tau) \in \Re^2 : s < t \le \tau \le f\}$, we may invoke Fubini's theorem for double integrals — see for example Apostol (1974), Theorem 15.6 — to conclude that

$$\int_s^f \lambda(t)\left(\int_t^f a(\tau)\,d\tau\right)dt = \iint\limits_{Z(s,f)} \lambda(t)a(\tau)\,d(t,\tau) = \int_s^f a(\tau)\left(\int_s^\tau \lambda(t)\,dt\right)d\tau, = \overline{C}(s,f)$$

□

A.2. Proof of Lemma 2

Proof.
$$\begin{aligned}\overline{C}(s,f) &= \int_s^v \lambda(t)\int_t^f a(u)\ du\ dt + \int_v^f \lambda(t)\int_t^f a(u)\ du\ dt \\ &= \int_s^v \lambda(t)\left(\int_t^v a(u)\ du + \int_v^f a(u)\ du\right) dt + \overline{C}(v,f) \\ &= \int_s^v \lambda(t)\int_t^v a(u)\ du\ dt + \int_s^v \lambda(t)\int_v^f a(u)\ du\ dt + \overline{C}(v,f) \\ &= \overline{C}(s,v) + \Lambda(s,v)A(v,f) + \overline{C}(v,f).\end{aligned}$$
□

A.3. Proof of Lemma 4

Proof. Consider first the case $s \leq f$. Noting that $\lambda(t) > 0$ and $a(\tau) \in [0, \bar{a}]$ in (2), we first obtain

$$0 \leq \overline{C}(s,f) \leq \int_s^f \lambda(t)\left(\int_t^f \bar{a}\ d\tau\right) dt = \int_s^f \lambda(t)\cdot\bar{a}\cdot(f-t)\ dt$$

Since $\lambda(t) \in [0, \bar{\lambda}]$ above, we further obtain

$$0 \leq \overline{C}(s,f) \leq \int_s^f \bar{\lambda}\bar{a}\cdot(f-t)\ dt = \frac{\bar{a}\bar{\lambda}}{2}(f-s)^2.$$

For the case $s > f$, we apply the same logic to $\int_f^s \lambda(t)\big(\int_f^t a(\tau)d\tau\big)dt$. □

A.4. Proof of Lemma 5

Proof. Applying Lemma 2 twice,

$$\begin{aligned}\overline{C}(s',f') &= \overline{C}(s',s) + \Lambda(s',s)A(s,f') + \overline{C}(s,f') \\ &= \overline{C}(s',s) + \Lambda(s',s)A(s,f') + \overline{C}(s,f) + \Lambda(s,f)A(f,f') + \overline{C}(f,f'),\end{aligned}$$

which rearranges to

$$\overline{C}(s',f') - \overline{C}(s,f) = \overline{C}(s',s) + \Lambda(s',s)A(s,f') + \Lambda(s,f)A(f,f') + \overline{C}(f,f').$$

Taking absolute values and applying the triangle inequality,

$$\left|\overline{C}(s,f) - \overline{C}(s',f')\right| \leq \overline{C}(s',s) + |\Lambda(s',s)|\,|A(s,f')| + |\Lambda(s,f)|\,|A(f,f')| + \overline{C}(f,f').$$

From this inequality, Lemma 4, the bounds on $a(\cdot)$ and $\lambda(\cdot)$, and $|s-s'|, |f-f'| \leq \epsilon$, we may then deduce

$$\begin{aligned}\left|\overline{C}(s,f) - \overline{C}(s',f')\right| &\leq \frac{\bar{a}\bar{\lambda}}{2}\epsilon^2 + \bar{\lambda}\epsilon\cdot\bar{a}\,|f'-s| + \bar{\lambda}\,|f-s|\cdot\bar{a}\epsilon + \frac{\bar{a}\bar{\lambda}}{2}\epsilon^2 \\ &\leq \bar{a}\bar{\lambda}\epsilon^2 + \bar{a}\bar{\lambda}\epsilon\,(|f-s|+\epsilon) + \bar{a}\bar{\lambda}\,|f-s|\,\epsilon \\ &= 2\bar{a}\bar{\lambda}\left(\epsilon^2 + |f-s|\,\epsilon\right).\end{aligned}$$
□

A.5. Proof of Proposition 6

Proof. We first note that Lemma 5 implies that $\overline{C}(\cdot,\cdot)$ is continuous on $[0,T]^2$, and the feasible region of (5) is closed and bounded. Therefore, the minimum value of (5) is attained by at least one feasible point $p^* = (p_0^*, p_1^*, \ldots, p_n^*) \in \Re^{n+1}$. Consider now the vector $\hat{p} = (\hat{p}_0, \hat{p}_1, \ldots, \hat{p}_n)$ obtained by rounding each p_i^* to the closest member of $\mathcal{T}$, rounding downwards in case of a tie. In both cases (i) and (ii), the maximum distance between elements of $\mathcal{T}$ is γ, so we have $|\hat{p}_i - p_i^*| \leq \gamma/2$ for all i. Applying Lemma 5, we have for all i that

$$\begin{aligned}\overline{C}(\hat{p}_{i-1}, \hat{p}_i) &\leq \overline{C}(p_{i-1}^*, p_i^*) + 2\bar{a}\bar{\lambda}\left((\gamma/2)^2 + (\gamma/2)(p_i^* - p_{i-1}^*)\right) \\ &= \overline{C}(p_{i-1}^*, p_i^*) + \bar{a}\bar{\lambda}\left(\gamma^2/2 + \gamma(p_i^* - p_{i-1}^*)\right)\end{aligned}$$

Adding the above inequality for $i = 1, \ldots, n$ produces

$$\begin{aligned}\overline{C}(\hat{p}_0, \hat{p}_1, \ldots, \hat{p}_n) &\leq \overline{C}(p_0^*, p_1^*, \ldots, p_n^*) + \bar{a}\bar{\lambda}\left(\sum_{i=1}^{n} \gamma^2/2 + \sum_{i=1}^{n} \gamma(p_i^* - p_{i-1}^*)\right) \\ &= C^* + \bar{a}\bar{\lambda}((n/2)\gamma^2 + T\gamma).\end{aligned}$$

Now, if $\hat{p}$ is feasible for (6), then $\ddot{C}$'s optimality for (6) implies $\ddot{C} \leq \overline{C}(\hat{p}_0, \hat{p}_1, \ldots, \hat{p}_n) \leq C^* + \bar{a}\bar{\lambda}((n/2)\gamma^2 + T\gamma)$. Therefore, it suffices to show that $\hat{p}$ is feasible for (6). By construction, $\hat{p}_i \in \mathcal{T}$ for all i, and since $t_0 = 0$ and $t_N = T$, we have $\hat{p}_0 = 0$ and $\hat{p}_n = t_N = T$. Thus, it is sufficient to show $\hat{p}_i - \hat{p}_{i-1} \geq \delta$ for $i = 1, \ldots, n$.

In case (i), we have $\delta = 0$, so we simply observe that since $p_{i-1}^* \leq p_i^*$ and the rounding operation is a nondecreasing map, we have $\hat{p}_{i-1} \leq \hat{p}_i$ for all i. Thus, it remains only to consider case (ii). We treat this remaining case by taking any $i \in \{1, \ldots, n\}$ and considering three subcases:

- First, suppose $\hat{p}_i \leq p_i^*$. Let $j \in \{1, \ldots, N\}$ be such that $\hat{p}_i = t_j$. Then

$$\begin{aligned} & p_i^* &\leq\ & \hat{p}_i + \gamma/2 \\ \Rightarrow\ & p_{i-1}^* &\leq\ & \hat{p}_i + \gamma/2 - \delta = t_{j-k} + \gamma/2 \\ \Rightarrow\ & \hat{p}_{i-1} &\leq\ & t_{j-k} = \hat{p}_i - \delta.\end{aligned}$$

- If we do not have $\hat{p}_i \leq p_i^*$, next suppose that $\hat{p}_{i-1} \geq p_{i-1}^*$. In this case, let $j \in \{0, \ldots, N\}$ be such that $t_j = \hat{p}_{i-1}$. Then

$$\begin{aligned} & p_{i-1}^* &\geq\ & \hat{p}_{i-1} - \gamma/2 \\ \Rightarrow\ & p_i^* &\geq\ & \hat{p}_{i-1} - \gamma/2 + \delta = t_{j+k} - \gamma/2 \\ \Rightarrow\ & \hat{p}_i &\geq\ & t_{j+k} = \hat{p}_{i-1} + \delta.\end{aligned}$$

- If neither of the above cases apply, then we must have $\hat{p}_i > p_i^*$ and $\hat{p}_{i-1} < p_{i-1}^*$. Then we simply deduce $\hat{p}_{i-1} < p_{i-1}^* \leq p_i^* - \delta < \hat{p}_i - \delta$. □

A.6. Enhancing Algorithm 4 to guarantee asymptotic local optimality

Unfortunately, $\overline{C}(\cdot,\cdot)$ and hence the objective function of (5) are in general nondifferentiable, raising the possibility of various kinds of "jamming" — that is, inability to make progress along any coordinate direction, even though the current point is not a local minimum. Even if we choose $\lambda(\cdot)$ and $a(\cdot)$ to be smoothly differentiable functions, making the objective function of (5) differentiable, the constraints $p_i - p_{i-1} \leq \delta$ can still induce a form of jamming. Such jamming could be overcome by applying techniques of constrained pattern search, as found for example in Kolda et al. (2003). In particular, one could adapt Algorithms 8.1 and 8.2 of Kolda et al. (2003) to produce schedule improvement methods with desirable asymptotic properties in the case that $\lambda(\cdot)$ and $a(\cdot)$ are smoothly differentiable. However, such methods would be far more complicated than Algorithm 4, because they would in general have to examine more search directions than Algorithm 4's simple axis-parallel directions. Specifically, the algorithm would have to identify "blocks" of consecutive probes that are within some distance $\epsilon > 0$ of being binding for the packing/sequencing constraints $p_i - p_{i-1} \leq \delta$. Such blocks would consist of indices $k, k+1, \ldots, k+\ell$ for which

$$\begin{array}{rcl} p_{k+1} & \leq & p_k + \delta + \epsilon \\ p_{k+2} & \leq & p_{k+1} + \delta + \epsilon \\ & \vdots & \\ p_{k+\ell} & \leq & p_{k+\ell-1} + \delta + \epsilon \end{array}$$

For each such block, the method would have to check 2ℓ possible search directions, corresponding to moving the first $m \leq \ell$ probes downward by the same offset or the last $m \leq \ell$ probes upward by the same offset, that is, for some $\theta > 0$ and $m \in \{1, \ldots, k\}$,

$$\begin{array}{lrclrcl} & p_i & \leftarrow & p_i - \theta & \forall\, i & = & k, k+1, \ldots, k+m \\ \text{or} & p_i & \leftarrow & p_i + \theta & \forall\, i & = & k+\ell-m, k+\ell-m+1, \ldots, k+\ell. \end{array}$$

Single coordinate moves, as practiced by Algorithm 4, are the $\ell = m = 0$ special case of this pattern.

In our view, however, the approximate global optimality properties of the discretized schedule obtained by Algorithm 1 are likely to be of far greater importance in practice than finding an exact *local* minimum of the expected cost function. Thus, we have not investigated such elaborate generalizations of Algorithm 4.

Additional References

Apostol, T. 1974. *Mathematical Analysis*. 2nd ed. Addison-Wesley.

Kolda, T.G., R.M. Lewis, V. Torczon. 2003. Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Rev.* **45** 385–482.