Electronic Companion: Proofs and Extensions

EC.1. Proof of Theorem 1

We introduce vector Lagrange multipliers $\alpha = (\alpha_1, \dots \alpha_{T-1})$ and $\beta = (\beta_1, \dots, \beta_{T-1})$, associated respectively with the capacity constraints $S_t \ge 0$ and $S_t \le E$ for $1 \le t \le T-1$, and a further multiplier μ_T associated with the constraint $S_T = S_T^*$. For any vector $S = (S_0, \dots, S_T)$ and for any α, β, μ_T as above, define the Lagrangian

$$L(S, \alpha, \beta, \mu_T) = \sum_{t=1}^{T} C_t(x_t(S)) - \sum_{t=1}^{T-1} [(\alpha_t + \beta_t) S_t - \beta_t E] - \mu_T (S_T - S_T^*). \tag{EC.1}$$

Suppose now that there exist some vectors S^*, α^* and β^* and some μ_T^* such that

(i') S^* is feasible for the problem **P**,

(ii') S^* minimises $L(S, \alpha^*, \beta^*, \mu_T^*)$ within the set of *all* S satisfying both $S_0 = S_0^*$ and the rate constraints (4),

(iii') S^*, α^* and β^* satisfy the complementary slackness conditions, for $1 \le t \le T-1$,

$$\begin{aligned} \alpha_t^* &\ge 0, \qquad \alpha_t^* = 0 \text{ when } S_t^* > 0, \\ \beta_t^* &\le 0, \qquad \beta_t^* = 0 \text{ when } S_t^* < E. \end{aligned}$$

Then, for any S satisfying all the constraints (3) and (4),

$$\sum_{t=1}^{T} C_t(x_t(S^*)) = L(S^*, \alpha^*, \beta^*, \mu_T^*) \le L(S, \alpha^*, \beta^*, \mu_T^*) \le \sum_{t=1}^{T} C_t(x_t(S)),$$

where the equality and the second inequality above follow from the above definition of the Lagrangian $L(S, \alpha, \beta, \mu_T)$ and the conditions (i'), (iii'), and where the first inequality follows from the condition (ii'). It thus follows that the vector S^* solves the problem **P**.

Now given a pair (S^*, μ^*) satisfying the conditions (i)–(iii) of the theorem, it follows from the condition (iii) that there exist (unique) vectors α^*, β^* satisfying the condition (iii') and such that

$$\mu_t^* = \rho^{T-t} \mu_T^* + \sum_{u=t}^{T-1} \rho^{u-t} (\alpha_u^* + \beta_u^*), \quad 1 \le t \le T-1. \tag{EC.2}$$

Further, from (1), $S_t = \rho^t S_0 + \sum_{u=1}^{t} \rho^{t-u} x_u(S)$ for $t = 1, \dots, T$ and for any S. It now follows from (EC.1) that for any vector S satisfying both $S_0 = S_0^*$ and the rate constraints (4),

$$\begin{aligned} L(S, \alpha^*, \beta^*, \mu_T^*) &= \sum_{t=1}^{T} C_t(x_t(S)) - \sum_{t=1}^{T-1} [(\alpha_t^* + \beta_t^*) S_t - \beta_t E] - \mu_T^* (S_T - S_T^*) \\ &= \sum_{t=1}^{T} C_t(x_t(S)) - \sum_{t=1}^{T-1} (\alpha_t^* + \beta_t^*) \sum_{u=1}^{t} \rho^{t-u} x_u(S) - \mu_T^* \sum_{u=1}^{T} \rho^{T-u} x_u(S) + k \\ &= \sum_{t=1}^{T} [C_t(x_t(S)) - \mu_t^* x_t(S)] + k, \end{aligned}$$

where k consists of terms which are constant over vectors S as above, and where the final inequality in the above display follows from (EC.2) on interchanging the roles of the subscripts t and u. Thus the condition (ii) of the theorem implies the condition (ii') above. Hence, finally, S^*, α^*, β^* and μ_T^* satisfy the conditions (i')–(iii') above, and so the vector S^* solves the problem $\mathbf{P}$. $\square$

EC.2. Proof of Theorem 2

Recall that the scalar $\hat{\mu}$ defined in Section 4 is such that $\mu_t^* = \rho^{1-t}\hat{\mu}$ and $S_t^* = S_t(\hat{\mu})$ for $1 \le t \le \tau$, where the functions $S_t(\cdot)$ are as given by (6) and where τ (the asserted first decision horizon) is also as defined in Section 4. It follows from the definitions of τ and $\hat{\mu}$ that $\mu^{l,t} \le \hat{\mu} \le \mu^{u,t}$ for $1 \le t \le \tau$; this follows since, when τ is defined via the condition (a) of Section 4, then $\hat{\mu} = \mu^{l,\tau} = \bar{\mu}^{l,\tau} < \bar{\mu}^{u,\tau}$ and so the claimed result follows from the definitions of $\bar{\mu}^{l,\tau}$ and $\bar{\mu}^{u,\tau}$; the argument when τ is defined via the condition (b) or the condition (c) is similar. It now follows from (8) that $a_t \le S_t^* \le b_t$ for $1 \le t \le \tau$, and so the constructed path $(S_1^*, \dots, S_\tau^*)$ (which by construction satisfies the rate constraints) is feasible for the problem $\mathbf{P}$ to the time τ. Hence, since the algorithm of Section 4 is restarted at the time τ and at subsequent similar times, it follows that the entire constructed path $(S_1^*, \dots, S_T^*)$ is feasible for the problem $\mathbf{P}$, and so the condition (i) of Theorem 1 is satisfied. Further, from the definitions in Section 4 of the functions $\hat{x}_t(\cdot)$ and $S_t(\cdot)$, and from those of μ_t^* and S_t^*, the constructed pair (S^*, μ^*) satisfies the condition (ii) of Theorem 1. Similarly, by construction the pair (S^*, μ^*) satisfies the complementary slackness conditions (iii) of Theorem 1, except perhaps at those times at which the algorithm of Section 4 is restarted.

In order to verify the condition (iii) at these remaining times, we make use of the following observation. Suppose that for some μ and for some $\bar{t} \le T$,

$$a_t < S_t(\mu) < b_t, \quad 1 \le t < \bar{t}, \qquad S_{\bar{t}}(\mu) \ge b_{\bar{t}}. \tag{EC.3}$$

Then necessarily $\hat{\mu} \le \mu$, where $\hat{\mu}$ is as defined by the algorithm of Section 4 (i.e. is such that $\mu_t^* = \rho^{1-t}\hat{\mu}$ and $S_t^* = S_t(\hat{\mu})$ for $1 \le t \le \tau$). To show this result we make use also of the fact that, from (8), the conditions (EC.3) are equivalent to

$$\bar{\mu}^{l,\bar{t}-1} < \mu < \bar{\mu}^{u,\bar{t}-1}, \qquad \mu \ge \mu^{u,\bar{t}}. \tag{EC.4}$$

The first of these relations implies that $\bar{\tau} \ge \bar{t}$. If $\bar{\tau} > \bar{t}$, then, regardless of which of the conditions (a)–(c) of Section 4 defines $\hat{\mu}$, we have $\hat{\mu} \le \bar{\mu}^{u,\bar{t}} \le \mu^{u,\bar{t}}$, so that the claimed result is here immediate on using the second relation in (EC.4). Suppose therefore that (EC.3) (or equivalently (EC.4)) holds and that $\bar{\tau} = \bar{t}$. We consider separately each of the three conditions (a)–(c) of Section 4. Under the condition (a) we have $\hat{\mu} = \bar{\mu}^{l,\bar{t}-1}$, so that, from (EC.4), $\hat{\mu} \le \mu$ as required. The condition (b) of Section 4 cannot hold here. To see this assume the contrary; the assumption $\bar{\tau} = \bar{t}$ implies that

$\tau < \bar{t}$; hence, from (EC.4), $\mu < \mu^{u,\tau}$ and, from (EC.3), $S_\tau(\mu) < b_\tau = S_\tau(\mu^{u,\tau})$; it now follows from (7) that

$$S_{\bar{t}}(\mu) < S_{\bar{t}}(\mu^{u,\tau}) \leq S_{\bar{t}}(\mu^{l,\bar{t}}) = a_{\bar{t}} \leq b_{\bar{t}},$$

where the second inequality above follows by the monotonicity of $S_{\bar{t}}(\cdot)$ and since, for $\bar{\tau} = \bar{t}$, the condition (b) implies $\mu^{l,\bar{t}} \geq \bar{\mu}^{u,\bar{t}-1} = \mu^{u,\tau}$; however, the relation $S_{\bar{t}}(\mu) < b_{\bar{t}}$ contradicts (EC.3). Finally, under the condition (c) of Section 4, we necessarily have $\bar{\tau} = T$ and so, for $\bar{\tau} = \bar{t}$, the conditions (EC.3) imply that $S_T(\mu) \geq S_T^*$; since here $\hat{\mu}$ is such that $S_T(\hat{\mu}) = S_T^*$, it follows from the monotonicity of $S_T(\cdot)$ that (or, in the event of nonuniqueness of $\hat{\mu}$, we may take) $\hat{\mu} \leq \mu$ as required.

Now suppose, without loss of generality, that the time τ (the asserted first decision horizon) and parameter $\hat{\mu}$ identified by the algorithm of Section 4 are determined via the condition (a) of that section. Since in this case we have $S_\tau^* = S_\tau(\mu^{l,\tau}) = a_\tau = 0$, in order to complete the proof it is necessary to show that $\rho\mu_{\tau+1}^* \leq \mu_\tau^*$. It follows from (13) that

$$a_t < S_t(\hat{\mu}) < b_t, \quad \tau < t < \bar{\tau}, \qquad\qquad S_{\bar{\tau}}(\hat{\mu}) \geq b_{\bar{\tau}},$$

and so, from the observation of the preceding paragraph, that $\hat{\mu}' \leq \rho^{-\tau}\hat{\mu}$ where $\hat{\mu}'$ plays the role of $\hat{\mu}$ in the algorithm restarted at the time τ (the factor $\rho^{-\tau}$ arising on account of role of ρ in the path definition (11)). Since $\mu_\tau^* = \rho^{1-\tau}\hat{\mu}$ and $\mu_{\tau+1}^* = \hat{\mu}'$, the required result now follows. □

EC.3. Relaxation of the strict convexity assumption for the algorithm

We discuss the modifications required to the algorithm of Section 4 when the convex cost functions C_t fail to be strictly convex. In practice non-strict convexity might be dealt with by some extremely small perturbation of these functions; however, a more formal approach may also be easily implemented. The problem here is that the functions $\hat{x}_t$ introduced in the description of the algorithm are no longer uniquely defined. Rather, for each time t and for each value of μ, $\hat{x}_t(\mu)$ may take any value in the closed interval $[x_t^l(\mu), x_t^u(\mu)]$, say, defining the set of minima x in X of the function $C_t(x) - \mu x$ (where we may of course have $x_t^l(\mu) = x_t^u(\mu)$). However, uniqueness may be restored by appending to the variable μ a second variable κ taking values in $[0,1]$ and, for each t, replacing $\hat{x}_t(\mu)$ by the well-defined function $\hat{x}_t(\mu,\kappa) = \kappa x_t^l(\mu) + (1-\kappa)x_t^u(\mu)$. If we now define a linear ordering on the set of possible values of (μ,κ) by $(\mu_1,\kappa_1) \leq (\mu_2,\kappa_2)$ if and only if either $\mu_1 < \mu_2$ or $\mu_1 = \mu_2$, $\kappa_1 \leq \kappa_2$, then (μ,κ) and $\hat{x}_t(\mu,\kappa)$ play the roles of μ and $\hat{x}_t(\mu)$ in the earlier theory. In particular, under the above linear ordering the function $\hat{x}_t(\mu,\kappa)$ is increasing as required, so that it is easily checked that the earlier theory goes through as before—removing in particular what would otherwise be some ambiguity in the definitions of the times $\bar{\tau}$ and τ.