Dynamic Portfolio Optimization with Transaction Costs: Heuristics and Dual Bounds

(Addendum on Gradient Penalties)

David B. Brown and James E. Smith

Fuqua School of Business
Duke University
Durham, NC 27708-0120

dbbrown@duke.edu, jes9@duke.edu

May 12, 2014

A. Addendum on Gradient Penalties

In this note, we discuss some of the "gradient penalties" used in Brown and Smith (2011), hereafter BS (2011), to calculate upper bounds for dynamic portfolio optimization problems with transaction costs. In particular, we discuss a simple way to improve some of these gradient penalties and present some updated numerical results for some of the numerical examples in BS (2011). This note is not intended to be fully self-contained and assumes the reader is familiar with BS (2011).

Building upon the general theory developed in Brown, Smith, and Sun (2010) and others, BS (2011) calculates upper bounds for dynamic portfolio optimization problems by solving a relaxed version of the problem in which the investor has perfect information on future returns, but has to pay a penalty for using this additional information. BS (2011) studies a number of penalties (and heuristics); one of the penalties is called the *frictionless gradient-based penalty* and is defined as

$$\hat{\pi}(\boldsymbol{a}, \boldsymbol{r}) \quad = \quad \nabla_{\boldsymbol{a}} U(w_T^f(\hat{\alpha}^*(\boldsymbol{r}), \boldsymbol{r}))'(\boldsymbol{a} - \alpha^*(\boldsymbol{r})). \tag{1}$$

Here, $\boldsymbol{a}$ denotes the trades (over all time periods) by the investor with perfect information, $\boldsymbol{r}$ denotes the vector of asset returns (over all time periods), U denotes the investor's utility function, w_T^f denotes the terminal (time T) wealth under the frictionless model (i.e., the return model without transaction costs), and $\hat{\alpha}^* = \hat{\alpha}^*(\boldsymbol{r})$ denotes the optimal trades in the frictionless model when returns are $\boldsymbol{r}$. BS (2011) also includes a random market state vector that may be correlated with returns; we omit that dependence here for simplicity. BS (2011) shows that penalties of the form (1) are dual feasible in that they lead to valid upper bounds and that these upper bounds will be no worse than the upper bound from the frictionless model itself. The dual feasibility result requires that the approximating model (in this case, the frictionless model) be solved to optimality.

Brown and Smith (2014), hereafter BS (2014), further develops the general theory of gradient penalties for *convex dynamic programs* (dynamic programs with convex feasible sets and concave reward functions). The gradient penalties discussed in BS (2014) are, like the penalty (1), linear in actions, but instead use gradients of approximate value functions, rather than of approximate reward functions. Applied to the portfolio optimization problem using the frictionless value function V_t^f as an approximate value function, these penalties have the form[1]

$$\hat{\pi}_{\triangledown}(\boldsymbol{a}, \boldsymbol{r}) \quad = \quad \sum_{t=0}^{T-1} \left(\nabla_{\boldsymbol{a}} V_{t+1}^f(\hat{\boldsymbol{\alpha}}_t^*) - \mathbb{E}\left[\nabla_{\boldsymbol{a}} V_{t+1}^f(\hat{\boldsymbol{\alpha}}_t^*) \;\middle|\; \mathcal{F}_t \right] \right)^{\top} (\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*). \tag{2}$$

In (2), $\boldsymbol{a}_t$ denotes the trades selected through time t in the problem with perfect information, $\hat{\boldsymbol{\alpha}}_t^*$ denotes the trades that are optimal for the frictionless model through time t, and $\mathcal{F}_t$ denotes the natural filtration in the original DP without additional information (i.e., returns and market state information up to time t).[2]

While studying these gradient penalties for BS (2014), we discovered that the "type 1" gradient penalties (1) and "type 2" gradient penalties (2) are related. In particular, it can be shown that the type 2 gradient penalties are better than the type 1 gradient penalties in every return scenario, i.e., $\hat{\pi}_{\triangledown}(\boldsymbol{a}, \boldsymbol{r}) \geq \hat{\pi}(\boldsymbol{a}, \boldsymbol{r})$ for all feasible actions $\boldsymbol{a}$, with equality holding along any return scenarios in which no constraints are binding in

[1] The gradient penalties in BS (2014) include zero mean terms that are independent of actions; these terms play the role of control variates and are important for variance reduction. BS (2011) uses these terms in the numerical examples but does not include these terms in the explicit definition of the penalty (1). We omit these terms here to simplify the notation.

[2] As in BS (2011), in (1) and (2) we take advantage of the fact that the frictionless wealth and frictionless value functions are differentiable; see BS (2014) for a treatment of the nondifferentiable case.

the frictionless model. Thus, the upper bounds using type 2 gradient penalties (2) will be "pathwise" better than the upper bounds using type 1 gradient penalties (1).

The intuition for this result is that, although the type 1 gradient penalties (1) include the effect of rewards, these penalties are potentially slack due to the omission of Lagrange multiplier terms that incorporate the impact of binding constraints. The type 2 gradient penalties capture the effect of such Lagrange multiplier terms implicitly through the use of continuation values. A formal proof of this result is provided at the end of this note; this proof is specialized to the portfolio optimization problem discussed in BS (2011) but extends to general convex DPs.

As stated, in order to use the type 1 gradient penalties (1), it is necessary that an approximating model (here, the frictionless model) be solved to optimality. Since this would usually require calculation of the associated approximating value functions as well, in practice we can get better upper bounds by using the type 2 gradient penalties (2) instead. These penalties do require the calculation of conditional expectations, although such expectations would usually need to be evaluated to solve the approximating model to optimality anyway.

In the numerical results in BS (2011), we reported a "frictionless gradient penalty" that uses a type 1 gradient penalty (1). Table 1 reports these results and shows a comparison to the upper bounds using the corresponding type 2 gradient penalty (2) for a set of examples from BS (2011). The upper bounds using the type 2 gradient penalties are always better than the upper bounds from the type 1 gradient penalties, as they must be. The differences are substantial (60% or more reduction in gap) in the examples with relatively low risk aversion: in these examples, the frictionless model invests heavily in a small number of assets, and thus the constraints (no short sales, i.e., weight of zero in some assets) are often binding. Conversely, with relatively high risk aversion, it is usually optimal in the frictionless model to invest in all of the assets (only in relatively low probability market states are some of the asset weights zero). Thus, in these cases the no short sales constraints are rarely binding and there is little difference in the penalties, and thus little difference in the upper bounds.

In summary, we recommend that other researchers using this approach to calculate upper bounds for convex dynamic programs rely on the "type 2" gradient penalties (2) whenever possible.

References

Brown, D.B., J.E. Smith. 2011. Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. *Management Science* 57 (10) 1752–1770.

Brown, D.B., J.E. Smith. 2014. Information relaxations, duality, and convex stochastic dynamic programs. Working paper, The Fuqua School of Business, Duke University.

Brown, D.B., J.E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Operations Research* 58 (4) 785–801.

Proof that $\hat{\pi}_\nabla(\boldsymbol{a},\boldsymbol{r}) \geq \hat{\pi}(\boldsymbol{a},\boldsymbol{r})$ for all feasible trades $\boldsymbol{a}$:

In order to show this result, we need to relate gradients of the frictionless value functions to gradients of the reward functions (in this case, the utility of the terminal wealth in the frictionless model). For a fixed set of trades $\boldsymbol{a}_{t\text{-}1}$ up to time t, we let $\boldsymbol{g}_t^f(\boldsymbol{a}_{t\text{-}1}, a_t)$ be the vector of constraint functions describing feasible trades a_t in period t in the frictionless model; $\boldsymbol{g}_t^f$ is concave in trades and implicitly depends on returns as well. Given past trades $\boldsymbol{a}_{t\text{-}1}$, we require that a_t be chosen such that $\boldsymbol{g}_t^f(\boldsymbol{a}_{t\text{-}1}, a_t) \geq \boldsymbol{0}$. Note that if we let $\boldsymbol{g}_t$ denote the vector of constraint functions in the model with transaction costs, then $\boldsymbol{g}_t^f(\boldsymbol{a}_t) \geq \boldsymbol{g}_t(\boldsymbol{a}_t)$, i.e., any trades that are feasible in the model with transaction costs are also feasible in the frictionless model.

We can write the frictionless value functions recursively; we take $V_T^f(\boldsymbol{a}) = U(w_T^f(\boldsymbol{a}))$ and for earlier times t have

$$V_t^f(\boldsymbol{a}_{t\text{-}1}) \quad = \max_{\left\{a_t \,:\, \boldsymbol{g}_t^f(\boldsymbol{a}_{t\text{-}1}, a_t) \geq \boldsymbol{0}\right\}} \mathbb{E}\Big[V_{t+1}^f(\boldsymbol{a}_{t\text{-}1}, a_t) \Big| \mathcal{F}_t \Big], \tag{3}$$

where expectations are taken over the next-period returns and market states. It is straightforward to show that V_t^f is concave in past trades. If strong duality holds in (3), then

$$V_t^f(\boldsymbol{a}_{t\text{-}1}) \quad = \quad \inf_{\boldsymbol{\lambda}_t \geq \boldsymbol{0}} \max_{a_t} \left\{ \boldsymbol{\lambda}_t^\top \boldsymbol{g}_t^f(\boldsymbol{a}_{t\text{-}1}, a_t) + \mathbb{E}\Big[V_{t+1}^f(\boldsymbol{a}_{t\text{-}1}, a_t) \Big| \mathcal{F}_t \Big] \right\}, \tag{4}$$

Assuming existence of an optimal Lagrange multiplier in (4), we can show that the frictionless value functions satisfy the recursion

$$\begin{pmatrix} \nabla V_t^f(\boldsymbol{a}_{t\text{-}1}) \\ \boldsymbol{0} \end{pmatrix} = \nabla \boldsymbol{\lambda}_t^\top \boldsymbol{g}_t^f(\boldsymbol{a}_{t-1}, a_t^*) + \mathbb{E}\Big[\nabla V_{t+1}^f(\boldsymbol{a}_{t\text{-}1}, a_t^*) \Big| \mathcal{F}_t \Big] \ , \tag{5}$$

where $\boldsymbol{\lambda}_t \geq \boldsymbol{0}$ are optimal Lagrange multipliers in (3) and a_t^* is an optimal trade in (3), given that past trades are $\boldsymbol{a}_{t\text{-}1}$.

Now consider a fixed set of trades $\boldsymbol{a} = \boldsymbol{a}_T$, feasible to the original portfolio problem (and thus also to the frictionless model), and a fixed set of returns $\boldsymbol{r}$. Let $\hat{\alpha}^*$ denote the optimal policy in the frictionless model.

We have

$$
\begin{aligned}
\hat{\pi}_{\triangledown}(\boldsymbol{a},\boldsymbol{r}) &= \sum_{t=0}^{T-1}\left(\nabla V_{t+1}^{f}(\hat{\boldsymbol{\alpha}}_t^*) - \mathbb{E}\Big[\nabla V_{t+1}^{f}(\hat{\boldsymbol{\alpha}}_t^*)\,\Big|\,\mathcal{F}_t\Big]\right)^{\top}(\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*) \\
&\overset{\text{(a)}}{=} \nabla V_T^{f}(\hat{\boldsymbol{\alpha}}^*)^{\top}(\boldsymbol{a} - \hat{\boldsymbol{\alpha}}^*) + \sum_{t=0}^{T-1}\left(\begin{pmatrix}\nabla V_t^{f}(\hat{\boldsymbol{\alpha}}_{t\text{-}1}^*) \\ \mathbf{0}\end{pmatrix} - \mathbb{E}\Big[\nabla V_{t+1}^{f}(\hat{\boldsymbol{\alpha}}_t^*)\,\Big|\,\mathcal{F}_t\Big]\right)^{\top}(\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*) \\
&\overset{\text{(b)}}{=} \nabla U(w_T^{f}(\hat{\boldsymbol{\alpha}}^*))^{\top}(\boldsymbol{a} - \hat{\boldsymbol{\alpha}}^*) + \sum_{t=0}^{T-1}\nabla\left(\boldsymbol{\lambda}_t^{*\top}\boldsymbol{g}_t^{f}(\hat{\boldsymbol{\alpha}}_t^*)\right)^{\top}(\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*) \\
&\overset{\text{(c)}}{=} \hat{\pi}(\boldsymbol{a},\boldsymbol{r}) + \sum_{t=0}^{T-1}\nabla\left(\boldsymbol{\lambda}_t^{*\top}\boldsymbol{g}_t^{f}(\hat{\boldsymbol{\alpha}}_t^*)\right)^{\top}(\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*) \\
&\overset{\text{(d)}}{\geq} \hat{\pi}(\boldsymbol{a},\boldsymbol{r}) + \sum_{t=0}^{T-1}\boldsymbol{\lambda}_t^{*\top}\left(\boldsymbol{g}_t^{f}(\boldsymbol{a}_t) - \boldsymbol{g}_t^{f}(\hat{\boldsymbol{\alpha}}_t^*)\right) \\
&\overset{\text{(e)}}{=} \hat{\pi}(\boldsymbol{a},\boldsymbol{r}) + \sum_{t=0}^{T-1}\boldsymbol{\lambda}_t^{*\top}\boldsymbol{g}_t^{f}(\boldsymbol{a}_t) \\
&\overset{\text{(f)}}{\geq} \hat{\pi}(\boldsymbol{a},\boldsymbol{r}) + \sum_{t=0}^{T-1}\boldsymbol{\lambda}_t^{*\top}\boldsymbol{g}_t(\boldsymbol{a}_t) \\
&\overset{\text{(g)}}{=} \hat{\pi}(\boldsymbol{a},\boldsymbol{r}).
\end{aligned}
$$

(a) follows by rearranging terms. (b) follows from the definition of V_T^f and by the recursion (5). (c) follows from the definition of $\hat{\pi}$, as given in (1). (d) follows from the fact that $\boldsymbol{\lambda}^{*\top}\boldsymbol{g}_t^f(\boldsymbol{a}_t)$ are concave functions (since $\boldsymbol{\lambda}_t^* \geq \mathbf{0}$ and $\boldsymbol{g}_t^f$ are concave) and hence

$$
\nabla\left(\boldsymbol{\lambda}_t^{*\top}\boldsymbol{g}_t^{f}(\hat{\boldsymbol{\alpha}}_t^*)\right)^{\top}(\boldsymbol{a}_t - \hat{\boldsymbol{\alpha}}_t^*) \quad \geq \quad \boldsymbol{\lambda}_t^{*\top}\left(\boldsymbol{g}_t^{f}(\boldsymbol{a}_t) - \boldsymbol{g}_t^{f}(\hat{\boldsymbol{\alpha}}_t^*)\right).
$$

(e) follows from the fact that $(\hat{\boldsymbol{\alpha}}_t^*, \boldsymbol{\lambda}_t^*)$ are primal-dual optimal to the frictionless model, so complementary slackness holds, i.e., $\lambda_{t,i}^* g_{t,i}^f(\hat{\boldsymbol{\alpha}}_t^*) = 0$ for each constraint i in each period t. (f) follows from the fact that $\boldsymbol{g}_t^f(\boldsymbol{a}_t) \geq \boldsymbol{g}_t(\boldsymbol{a}_t)$, and (g) follows from the fact that $\boldsymbol{a}$ is a feasible trade, and thus $\boldsymbol{g}_t(\boldsymbol{a}_t) \geq \mathbf{0}$.

Note that if the constraints are not binding in the frictionless model in this return scenario, then complementary slackness requires $\boldsymbol{\lambda}_t^* = \mathbf{0}$, and equality holds above throughout; in this case, the penalties are equal. □

Table 1: Results for the Three-Asset Model with Predictability

Parameters				Lower Bound	Upper Bounds			Gap Improvement		
orizon (T)	Risk Aversion Coeff. (γ)	Trans. Cost Rate (δ)		Best Heuristic Strategy	Frictionless Gradient Penalty (Type 2)	Frictionless Gradient Penalty (Type 1)	No Trans. Cost Bound	Type 2 Gap	Type 1 Gap	Gap Reduction (%)
12	1.5	0.5%	CE Return (%)	6.57	6.70	6.95	7.25	0.13	0.38	65.2%
			Mean Std. Error (%)	0.13	0.00	0.01				
			Turnover (%)	9.5	7.7	4.5	42.3			
12	1.5	1.0%	CE Return (%)	6.01	6.25	6.69	7.25	0.24	0.68	64.5%
			Mean Std. Error (%)	0.15	0.01	0.01				
			Turnover (%)	8.3	6.8	4.1	42.3			
12	1.5	2.0%	CE Return (%)	5.03	5.51	6.23	7.25	0.48	1.20	59.7%
			Mean Std. Error (%)	0.19	0.02	0.02				
			Turnover (%)	6.6	5.3	3.5	42.3			
12	3	0.5%	CE Return (%)	3.50	3.81	3.82	3.99	0.31	0.31	1.7%
			Mean Std. Error (%)	0.07	0.00	0.00				
			Turnover (%)	6.0	2.7	2.7	21.3			
12	3	1.0%	CE Return (%)	3.21	3.65	3.66	3.99	0.44	0.45	1.4%
			Mean Std. Error (%)	0.10	0.01	0.01				
			Turnover (%)	4.7	2.5	2.5	21.3			
12	3	2.0%	CE Return (%)	2.69	3.38	3.39	3.99	0.69	0.70	0.9%
			Mean Std. Error (%)	0.12	0.01	0.01				
			Turnover (%)	3.4	2.2	2.2	21.3			
12	8	0.5%	CE Return (%)	1.60	1.73	1.73	1.79	0.13	0.13	1.5%
			Mean Std. Error (%)	0.03	0.00	0.00				
			Turnover (%)	2.3	1.0	1.0	8.1			
12	8	1.0%	CE Return (%)	1.50	1.67	1.67	1.79	0.17	0.18	1.3%
			Mean Std. Error (%)	0.04	0.00	0.00				
			Turnover (%)	1.8	0.9	0.9	8.1			
12	8	2.0%	CE Return (%)	1.31	1.57	1.57	1.79	0.26	0.26	0.8%
			Mean Std. Error (%)	0.05	0.01	0.01				
			Turnover (%)	1.3	0.8	0.8	8.1			