Trading on a Mean-Reverting Asset
via Worst-Case Value-at-Risk Maximization

Alexandre d'Aspremont and Laurent El Ghaoui

September 2010

1 Investment problem

Asset and cash processes. We consider two univariate discrete-time processes. One called the "asset process" described by

$$x(t+1) - m = \lambda(x(t) - m) + \sigma w(t), \;\; w(t) \sim \mathcal{N}(0, 1),$$

where m, λ and $\sigma > 0$ and the initial asset price $x(0)$ are given, and $\lambda > 0$ is a parameter that enforces mean reversion.

In addition we have the "cash process" described by

$$B(t) = \mu^t B(0), \;\; t = 0, 1, 2, \ldots,$$

where $\mu = 1 + r$, with $r > 0$ the interest rate of cash, and $B(0)$ is the initial cash value.

Profit process. The investment problem over horizon T consists in starting with a budget $P(0)$ at time $t = 0$ and investing at each time $t = 0, \ldots, T$ a certain proportion in the risky asset relative to cash. The portfolio's value, or "profit process", is described by

$$P(t) = \alpha(t)x(t) + \beta(t)B(t),$$

where $\alpha(t)$ is the number of stocks. We enforce the self-financing condition

$$P(t+1) = P(t) + \alpha(t)(x(t+1) - x(t)) + \beta(t)rB(t), \;\; t = 0, \ldots, T.$$

Budget constraint. In addition we have budget constraints of the form

$$P(t) \geq L, \;\; t = 1, \ldots, T-1,$$

where $L > 0$ is a lower bound on the portfolio value. Note that the constraint involving the final portfolio value $P(T)$ is omitted as we will treat it separately, as seen next. We enforce the budget constraints in probability:

$$\text{Prob } \{P(t) < L\} \leq \epsilon, \;\; t = 1, \ldots, T-1,$$

where $\epsilon \in (0, 1)$ is a parameter.

Value-at-risk objective. To measure performance we consider a modified budget constraint for the final value $P(T)$, of the form

$$\text{Prob } \{P(T) < L(T)\} \leq \epsilon, \;\; t = 1, \ldots, T,$$

and will seek to maximize $L(T)$, with the lower bound $L(T) \geq L$.

Summary. To summarize: our investment problem is modeled as: Given $P(0)$ and $\alpha(0)$, $B(0)$, solve, for $u := (\alpha(0), \ldots, \alpha(T))$:

$$\begin{array}{ll}
\max\limits_{L(T),u} & L(T) - c \cdot \mathbf{E}_w \sum\limits_{t=0}^{T-1} |\alpha(t+1) - \alpha(t)| \\
\text{s.t.} & P(t) = \alpha(t)x(t) + \beta(t)B(t) \;\; t = 0, \ldots, T-1, \\
& \text{Prob } \{P(t) < L\} \le \epsilon, \;\; t = 1, \ldots, T-1, \\
& \text{Prob } \{P(T) < L(T)\} \le \epsilon, \;\; L(T) \ge L, \\
& P(t+1) = P(t) + \alpha(t)(x(t+1) - x(t)) + \beta(t)rB(t), \;\; t = 0, \ldots, T-1 \\
& x(t+1) = m + \lambda(x(t) - m) + \sigma w(t), \;\; B(t+1) = \mu B(t), \;\; t = 0, \ldots, T-1 \\
& u_1 = \alpha(0).
\end{array}$$

2 Profit as Bilinear Function

In this section our goal is to express the profit $P(t)$ as a bilinear function of the control variable u and noise vector w.

Vector form of asset process. We can express the asset process an any point in time as a (strictly causal) affine function of $w := (w(0), \ldots, w(T-1))$ using matrix notation, as follows. With $\delta(t) := x(t) - m$, we have

$$\delta(t+1) = \lambda\delta(t) + \sigma w(t), \;\; t = 0, \ldots, T-1.$$

In vector form:

$$(I - \lambda Z)\delta = \lambda\delta(0)e_1 + \sigma w,$$

with $\delta := (\delta(1), \ldots, \delta(T)) \in \mathbf{R}^T$, e_1 the first unit vector in $\mathbf{R}^T$, and Z the $T \times T$ backward shift matrix:

$$Z = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 1 & 0 \end{pmatrix}.$$

The above leads to

$$\delta = \delta^0 + \sigma(I - \lambda Z)^{-1}w, \;\; \delta^0 := \lambda\delta(0)(I - \lambda Z)^{-1}e_1.$$

The value process. Expressing $\beta(t)$ as a function of $\alpha(t)$ with the first set of equality constraints:

$$\beta(t)B(t) = P(t) - \alpha(t)x(t), \;\; t = 0, \ldots, T,$$

we obtain the recursion for the value process P:

$$P(t+1) = \mu P(t) + d(t)\alpha(t), \;\; t = 0, \ldots, T-1.$$

where

$$d(t) := x(t+1) - \mu x(t) = \delta(t+1) - \mu\delta(t) - rm, \;\; t = 0, \ldots, T-1.$$

With $P = (P(1), \ldots, P(T))$, and $u = (\alpha(0), \ldots, \alpha(T))$, we can write:

$$(I - \mu Z)P = \mu P(0)e_1 + \mathbf{diag}(Ru)d,$$

where $R = [I_T, 0]$ is a $T \times (T+1)$ projection matrix, such that $Ru = (\alpha(0), \ldots, \alpha(T-1))$.

Expression of d. Let us express the process d in vector form:

$$d := \left(\begin{array}{c} d(0) \\ \vdots \\ d(T-1) \end{array}\right) = \delta - \mu \left(\begin{array}{c} \delta(0) \\ \vdots \\ \delta(T-1) \end{array}\right) - rm \cdot \mathbf{1},$$

with $\mathbf{1}$ the vector of ones in $\mathbf{R}^T$. The vector d reads

$$d = (I - \mu Z)\delta - \mu\delta(0)e_1 - rm \cdot \mathbf{1}.$$

Replacing δ with its expression as a function of w we obtain

$$d = d^0 + Dw,$$

where

$$D = \sigma(I - \mu Z)(I - \lambda Z)^{-1}$$

and

$$\begin{array}{rcl} d^0 & = & (I - \mu Z)\delta^0 - \mu\delta(0)e_1 - rm \cdot \mathbf{1} \\ & = & \left(\lambda(I - \mu Z)(I - \lambda Z)^{-1} - \mu I\right)\delta(0)e_1 - rm \cdot \mathbf{1} \\ & = & \left(\lambda(I - \mu Z) - \mu(I - \lambda Z)\right)(I - \lambda Z)^{-1}\delta(0)e_1 - rm \cdot \mathbf{1} \\ & = & (1 - \dfrac{\mu}{\lambda})\delta^0 - rm \cdot \mathbf{1}. \end{array}$$

Process P as a bilinear function. We can then write $P(t)$ as a bilinear function in $u = (\alpha(0), \ldots, \alpha(T))$, and w. With $P = (P(1), \ldots, P(T))$, and $d = d^0 + Dw$, we have

$$\begin{array}{rcl} (I - \mu Z)P & = & \mu P(0)e_1 + \mathbf{diag}(Ru)d \\ & = & \mu P(0)e_1 + \mathbf{diag}(Ru)d^0 + \mathbf{diag}(Ru)Dw. \end{array}$$

Thus, using the fact that $\mathbf{diag}(u)v = \mathbf{diag}(v)u$ for any two compatible vectors,

$$P \;=\; P^0 + Au + (I - \mu Z)^{-1}\,\mathbf{diag}(Ru)Dw,$$

with

$$P^0 := \mu P(0)(I - \mu Z)^{-1}e_1, \;\; A := (I - \mu Z)^{-1}\,\mathbf{diag}(d^0)R = [(I - \mu Z)^{-1}\,\mathbf{diag}(d^0), 0].$$

We obtain, with e_t the t-th unit vector in $\mathbf{R}^T$, $t = 1, \ldots, T$:

$$\begin{array}{rcl} P(t) = e_t^T P & = & e_t^T P^0 + e_t^T Au + e_t^T (I - \mu Z)^{-1}\,\mathbf{diag}(Ru)Dw \\ & = & e_t^T P^0 + e_t^T Au + u^T R^T\,\mathbf{diag}((I - \mu Z)^{-T}e_t)Dw \\ & = & p(t) + u^T(Q(t)w + q(t)), \;\; t = 1, \ldots, T, \end{array}$$

where

$$Q(t) := R^T\,\mathbf{diag}((I - \mu Z)^{-T}e_t)D, \;\; q(t) := A^T e_t, \;\; p(t) := e_t^T P^0, \;\; t = 1, \ldots, T.$$

Here, matrix $Q(t)$, vector $q(t)$ and scalar $p(t)$ contain problem parameters, such as m, λ. In Appendix A, we provide the above data as functions of the original parameters of the investment problem.

3 Simple Stochastic Counterparts

Nominal problem. If w was known, we could write the problem of maximizing the terminal profit $P(T)$ in abstract form as

$$\begin{array}{ll} \max\limits_{u\,:\,u_1=\alpha(0)} & u^T Q(T) w + q(T)^T u - c\|Bu\|_1 \\ \text{s.t.} & u^T Q(t) w + q(t)^T u + p(t) \geq L, \;\; t = 1, \ldots, T, \end{array}$$

where the bi-diagonal matrix $B \in \mathbf{R}^{T \times (T+1)}$ reflects the transaction costs:

$$B = \left(\begin{array}{ccccc} -1 & 1 & 0 & \ldots & 0 \\ 0 & -1 & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & -1 & 1 \end{array} \right).$$

The above optimization problem is an LP, with $O(T)$ variables and constraints.

Stochastic counterpart. When w is not known, a first approach is to consider a stochastic counterpart of the above LP, with no recourse (that is, u is to be decided at the first period and not modified thereafter). When $w \sim \mathcal{N}(0, I)$ is Gaussian, and we seek to minimize the Value-at-Risk, we end up with a problem of the form

$$\begin{array}{ll} \max\limits_{u\,:\,u_1=\alpha(0)} & q(T)^T u - \kappa(\epsilon)\|Q(T)^T u\|_2 - c\|Bu\|_1 \\ \text{s.t.} & q(t)^T u + p(t) \geq L + \kappa(\epsilon)\|Q(t)^T u\|_2, \;\; t = 1, \ldots, T, \end{array}$$

where $\kappa(\epsilon) = -\Phi^{-1}(\epsilon)$, with Φ the cumulative density function of the normal distribution. The above is an SOCP.

Distributionnally robust approach. We may also decide not to assume that w is Gaussian, and instead assume that all is known about w is that its mean is zero and its covariance matrix is the identity. In this case we can replace the probability constraints, for example

$$\text{Prob } \{P(t) < L\} \leq \epsilon$$

with the "distributionally robust" counterpart

$$\sup_{w \sim (0,I)} \text{ Prob } \{P(t) < L\} \leq \epsilon$$

where the sup is taken with respect to all the distributions of w with zero mean and identity covariance matrix. In this case the problem takes the same form as in the previous Gaussian case, with $\kappa(\epsilon) = \sqrt{(1-\epsilon)/\epsilon}$.

4 Affine Recourse Approach

In a affine recourse approach we set the control variable to be a strictly causal, affine function of the uncertainty. The coefficients of the affine function become the new decision variables. Specifically, we replace u with the strictly causal affine function $u + Uw$, where u, U are both variables, with U a $(T+1) \times T$ strictly lower-triangular matrix.

4.1 Profit process

The profit variable $P(t)$, $t = 1, \ldots, T$ becomes a bilinear function of u, U and w:

$$P(t) = p(t) + (u + Uw)^T (Q(t)w + q(t)) = \left(\begin{array}{c} w \\ 1 \end{array} \right)^T \Pi(U, u, t) \left(\begin{array}{c} w \\ 1 \end{array} \right), \quad t = 1, \ldots, T$$

where

$$\Pi(U, u, t) \quad := \quad \frac{1}{2} \left(\begin{array}{cc} U^T Q(t) + Q(t)^T U & U^T q(t) + Q(t)^T u \\ q(t)^T U + u^T Q(t) & 2(p(t) + q(t)^T u) \end{array} \right), \quad t = 1, \ldots, T.$$

4.2 Budget Constraints

For given $t \in \{1, \ldots, T\}$, the budget constraint

$$\text{Prob } \{P(t) < L\} \leq \epsilon$$

is equivalent to one of the form

$$\text{Prob } \left\{ (w, 1)^T W (w, 1) > 0 \right\} \leq \epsilon$$

with $W := L \cdot J - \Pi(U, u, t)$, where J is a $(T + 1) \times (T + 1)$ matrix having zeros everywhere, except a 1 at the bottom right entry. For now, u, U and t are fixed so we omit the (affine) dependence of W on these variables.

Distributionally robust approach. Even when w is restriced to be Gaussian, there does not seem to be a closed-form expression for the probability above. Hence we take a worst-case approach and consider the problem of checking if

$$\sup_{w \sim (0, I)} \mathbf{Prob}(x \in \mathcal{W}) \leq \epsilon$$

holds, where the $\sup$ is taken with respect to all distributions having zero mean and unit covariance matrix, and $\mathcal{W}$ is given by

$$\mathcal{W} = \left\{ w \ : \ (w, 1)^T W (w, 1) > 0 \right\}.$$

By duality,

$$P_{\text{wc}} = \min_{M = M^T} \mathbf{Tr}\, M \ : \ \begin{array}{l} \forall\, w \in \mathcal{W}, \ \ (w, 1)^T M (w, 1) \geq 1, \\ \forall\, w \notin \mathcal{W}, \ \ (w, 1)^T M (w, 1) \geq 0, \end{array}$$

which reduces to the equivalent problem

$$P_{\text{wc}} = \min_{M = M^T \succeq 0} \mathbf{Tr}\, M \ : \ \forall\, w \in \mathcal{W}, \ \ (w, 1)^T M (w, 1) \geq 1.$$

For given $M = M^T$, the condition

$$\forall\, w \in \mathcal{W}, \ \ (w, 1)^T M (w, 1) \geq 1$$

is equivalent to the existence of a scalar $\eta \geq 0$ such that

$$\forall\, w \ : \ (w, 1)^T M (w, 1) \geq 1 + \eta (w, 1)^T W (w, 1),$$

that is: $M \succeq J + \eta W$. We obtain

$$P_{\text{wc}} = \min_{M, v} \mathbf{Tr}\, M \ : \ M \succeq 0, \ \ \eta \geq 0, \ \ M \succeq J + \eta W$$

The condition $P_{\rm wc} \le \epsilon$ is thus equivalent to the LMI in $M, v := 1/\eta$:

$$\mathbf{Tr}\, M \le \epsilon v, \quad M \succeq 0, \quad M \succeq vJ + W.$$

Note that the first constraint ensures that $v \ge 0$.

Our (conservative) probabilistic budget constraint requires the existence of scalars $v(t)$, and matrix $M(t)$ such that

$$\mathbf{Tr}\, M(t) \le \epsilon v(t), \quad M(t) \succeq 0, \quad M(t) + \Pi(U,u,t) \succeq (L + v(t))J.$$

The above is an LMI in U, u and $M(t), v(t)$.

4.3 Expected transaction costs

Worst-case approach. The expected transaction costs are $cF(U,u)$, with

$$F(U,u) := \mathbf{E}_w \, \|S(U)w + s(u)\|_1,$$

where $S(U) := BU \in \mathbf{R}^{T\times T}$, $s(u) := Bu \in \mathbf{R}^T$ are linear in u. Here again we take a worst-case approach and consider the quantity

$$F := \sup_{w\sim(0,I)} \mathbf{E}_w \, \|Sw + s\|_1,$$

where $S \in \mathbf{R}^{T\times T}$, $s \in \mathbf{R}^T$ are given.

Intermediate result. We can bound F as follows. For $t = 1, \ldots, T$, denote by e_t the t-th unit vector in $\mathbf{R}^T$, and by $S_t^T = e_t^T S$ the t-th row of S. For any vector random variable $\xi \in \mathbf{R}^T$, and using the concavity of the square root function:

$$\mathbf{E}_\xi \, \|\xi\|_1 = \mathbf{E}_\xi \, \sum_{t=1}^T \sqrt{\xi_t^2} \le \sum_{t=1}^T \sqrt{\mathbf{E}_\xi \, \xi_t^2}.$$

Applying this to $\xi = Sw + s$, we obtain

$$\begin{array}{rcl} F = \mathbf{E}_w \, \|\xi\|_1 & \le & \psi := \displaystyle\sum_{t=1}^T \sqrt{\mathbf{E}\, \xi_t^2} \\ & = & \displaystyle\sum_{t=1}^T \sqrt{\mathbf{E}(S_t^T w + s_t)^2} \\ & = & \displaystyle\sum_{t=1}^T \sqrt{\|S_t\|_2^2 + s_t^2} \\ & = & \displaystyle\sum_{t=1}^T \left\| \left(\begin{array}{c} S_t \\ s_t \end{array} \right) \right\|_2 . \end{array}$$

It turns out that the bound is tight within a $\pi/2$ factor, that is, $F \ge (2/\pi)\psi$.

Bound on expected transaction costs. Let us use our intermediate result, with $s_t = e_t^T Bu = u_{t+1} - u_t$, and

$$S_t^T = e_t^T BU = (e_{t+1} - e_t)^T U = U(t+1,\cdot) - U(t,\cdot), \quad t = 1, \ldots, T,$$

where, for $t = 1, \ldots, T+1$, $U(t,\cdot)$ is the t-th row of matrix U. The expected transaction costs are then bounded below by

$$c \cdot F(U,u) \le c \cdot \sum_{t=1}^T \left\| \left(\begin{array}{c} (U(t+1,\cdot) - U(t,\cdot))^T \\ u_{t+1} - u_t \end{array} \right) \right\|_2 .$$

4.4 Summary

Problem statement. To summarize, our approximation to the investment problem takes the form of an SDP-representable problem:

$$\begin{array}{rl} \max\limits_{u,U,L(T),M(t),v(t),t=1,\ldots,T} & L(T) - c \cdot \sum_{t=1}^{T} \left\| \begin{pmatrix} (U(t+1,\cdot) - U(t,\cdot))^T \\ u_{t+1} - u_t \end{pmatrix} \right\|_2 \\ \text{s.t.} & \mathbf{Tr}\, M(t) \leq \epsilon v(t), \ \ M(t) \succeq 0, \ \ t = 1, \ldots, T, \\ & M(t) + \Pi(U,u,t) \succeq (L(t) + v(t))J, \ \ t = 1, \ldots, T, \\ & L(t) = L, \ \ t = 1, \ldots, T-1, \ \ L(T) \geq L, \\ & u_1 = \alpha(0). \end{array}$$

The above problem may not be feasible. To fix this we can let $L(t)$, $t = 1, \ldots, T-1$, become variables and solve the problem

$$\begin{array}{rl} \max\limits_{u,U,L(t),M(t),v(t),t=1,\ldots,T} & L(T) - b \cdot \sum_{t=1}^{T} (L - L(t))_+ - c \cdot \sum_{t=1}^{T} \left\| \begin{pmatrix} (U(t+1,\cdot) - U(t,\cdot))^T \\ u_{t+1} - u_t \end{pmatrix} \right\|_2 \\ \text{s.t.} & \mathbf{Tr}\, M(t) \leq \epsilon v(t), \ \ M(t) \succeq 0, \ \ t = 1, \ldots, T, \\ & M(t) + \Pi(U,u,t) \succeq (L(t) + v(t))J, \ \ t = 1, \ldots, T, \\ & u_1 = \alpha(0), \end{array}$$

where $b > 0$ is a parameter.

A Problem data

Let us summarize the dependence of the problem's data on the investment and model parameters, as given in Table 1.

T :	investment horizon.
λ :	rate of mean reversion.
r :	rate of cash.
m :	asset mean.
σ :	standard deviation of noise.
$x(0)$:	initial asset price.
$P(0)$:	initial budget.
$\alpha(0)$:	initial position in asset.
$B(0)$:	initial cash value.
c :	rate of linear transaction costs.
L :	lower bound on profit.
ϵ :	bound on probability of profit to be below level L.

Table 1: Table of problem parameters.

We define the matrices

$$R = \begin{pmatrix} I_T & 0_{T\times 1} \end{pmatrix} \in \mathbf{R}^{T\times(T+1)}, \ \ Z = \begin{pmatrix} 0_{1\times(T-1)} & 0_{1\times 1} \\ I_{T-1} & 0_{(T-1)\times 1} \end{pmatrix} \in \mathbf{R}^{T\times T},$$

and

$$B = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 1 \end{pmatrix} \in \mathbf{R}^{T\times(T+1)}, \;\; J = \mathbf{diag}(0_{T\times T}, 1) \in \mathbf{R}^{(T+1)\times(T+1)}.$$

The problem's data are then functions of the initial parameters that are given in Table 1, as follows:

$$\begin{array}{rcl} \mu & = & 1 + r, \\ G & = & (I - \lambda Z)^{-1}, \\ H & = & (I - \mu Z)^{-1}, \\ \delta^0 & = & \lambda(x(0) - m)Ge_1, \\ d^0 & = & (1 - \frac{\mu}{\lambda})\delta^0 - rm \cdot \mathbf{1}, \\ D & = & \sigma(I - \mu Z)G, \\ P^0 & = & \mu P(0)He_1, \\ A & = & H\,\mathbf{diag}(d^0)R, \\ Q(t) & = & R^T\,\mathbf{diag}(H^T e_t)D, \;\; t = 1, \dots, T, \\ q(t) & = & A^T e_t, \;\; t = 1, \dots, T, \\ p(t) & = & e_t^T P^0, \;\; t = 1, \dots, T, \\ \Pi(U, u, t) & = & \frac{1}{2}\begin{pmatrix} U^T Q(t) + Q(t)^T U & U^T q(t) + Q(t)^T u \\ q(t)^T U + u^T Q(t) & 2(p(t) + q(t)^T u) \end{pmatrix}, \;\; t = 1, \dots, T. \end{array}$$