Nonlinear Optimal Trade-Off Control for LQG Problem

Fucai Qian, Guo Xie, Ding Liu and Wenfang Xie

Abstract—For discrete-time Linear Quadratic Gaussian (LQG) control problems, a utility function on the expectation and the variance of the conventional performance index, is considered. The utility function is viewed as an overall objective of system. The nonlinear utility function is first converted into an auxiliary parameters optimization problem about the expectation and the variance. Then an optimal closed-loop feedback controller for the nonseparable mean-variance minimization problem is designed by nonlinear mathematical programming. Finally, simulation results are given to verify the algorithm’s effectiveness obtained in this paper.

I. INTRODUCTION

It is well-known that linear-quadratic Gaussian control [2] represents one of the most prominent successes in control theory, largely due to its wide applications and its mathematical elegance in tractability. In the last forty years, the study of LQG has made great achievement covering risk sensitive optimal control [20], multiple quadratic performance indices[3], [5], [11], [12], mean-variance control of the performance index [7], [8], [14], cumulant control of the performance index [15], [16] and cost smoothing of the performance index over time [10]. The common feature of these above control paradigms is that the overall objective function is the linear combinations of the expected value $E(J)$ and the variance $Var(J)$ under mean-variance framework or the linear combination of the first $k$ cost cumulants under cumulant control and Risk-Sensitive control framework, where $J$ is the quadratic performance index of LQG problem. However, in many situations, the overall objective functions are the nonlinear functions of $E(J)$ and $Var(J)$.

For example, in investment, the investor can determine a portfolio by maximizing some utility functions such as $U(x) = 1 - e^{-bx}$ ($b > 0$), $U(x) = x^a$ ($0 < a < 1$) and $U(x) = \log x$ where $x$ is the investor’s return. Such utility functions satisfy the condition that their second derivative is a nondecreasing function, which means $U''(x)$ is nondecreasing in $x$. Let $x = J$ and $\mu = E(J)$ representing the performance of investment. One can approximate $U(J)$ by using the first three terms of its Taylor series expansion about the point $\mu$. That is, we use the approximation

$$U(J) \approx U(\mu) + U'(\mu)(J - \mu) + \frac{1}{2} U''(\mu)(J - \mu)^2.$$ 

Taking expectations gives that

$$E[U(J)] \approx U(\mu) + U'(\mu)E[J - \mu] + \frac{1}{2} U''(\mu)E[(J - \mu)^2]$$

$$= U(\mu) + \frac{1}{2} U''(\mu)Var^2(J).$$

Therefore, a reasonable approximation to the optimal portfolio is given by the portfolio that maximizes

$$U(E(J)) + \frac{1}{2} U''(E(J))Var^2(J). \quad (1)$$

It is clear that the expression (1) is a nonlinear function of the expected value and the variance of $J$.

In the recent developed performance first control for LQG problems, the objective function is taken as $P(J \geq J)$ in the minimization sense, where $J$ is a so-called disaster level and is assumed to be greater than $E(J)$ associated with the control policy adopted and is the ruin probability. From the Bienayme-Tchebycheff inequality, one has

$$P(J \geq J) = \mathbb{P}[(J - E(J)) \geq J - E(J)] \leq \frac{Var(J)}{(E(J) - J)^2}. \quad (2)$$

As suggested by Roy’s safety-first principle (Roy 1952), minimizing $P(J \geq J)$ can be achieved by minimizing its upper bound. Therefore, one may consider the surrogate problem $\min G \{E(J), Var(J)\} = \frac{Var(J)}{(E(J) - J)^2}$ to solve the performance first minimization problem. Note that objective function $G$ is also a nonlinear function of two evaluation criteria, $E(J)$ and $Var(J)$. While an optimal control law is developed specifically for such nonlinear function, the purpose of this paper is to derive an optimal closed-loop control law for a more general nonlinear objective function which is convex.

II. PROBLEM DESCRIPTION

We consider in this paper the following discrete-time linear-quadratic Gaussian optimal control problem,

$$z(k + 1) = A(k)z(k) + B(k)u(k) + \xi(k) \quad (3)$$

$$k = 0, 1, \ldots, N - 1,$$

where $z(k) \in R^n$ is the state, $u(k) \in R^m$ is the control, $A(k)$ and $B(k)$ are matrices of appropriate dimensions, $\{\xi(k)\}$ is a sequence of white Gaussian random noises with $\xi(k) \sim N(0, \Theta_\xi)$ and the initial state $z(0)$ is known.
The performance index of system (3) is the quadratic function of state and control, i.e.,
\begin{align}
J &= z^T(N)Q(N)z(N) \\
&+ \sum_{k=0}^{N-1} [z^T(k)Q(k)z(k) + u^T(k)R(k)u(k)],
\end{align}
where it is assumed that for all \(k\) the matrices \(Q(k)\) and \(R(k)\) are symmetric positive semidefinite and positive definite, respectively.

The optimal control for the traditional linear-quadratic Gaussian control problem of \(\min\{E(J) \mid \text{subject to (3)}\}\) is well known as
\[u^*(k) = -L(k)z(k)\]
where
\begin{align*}
L(k) &= D^{-1}(k)B^T(k)S(k+1)A(k) \\
D(k) &= B^T(k)S(k+1)B(k) + R(k) \\
S(k) &= A^T(k)S(k+1)A(k) - L^T(k)L(k) + Q(k)
\end{align*}
with boundary condition of \(S(N) = Q(N)\). The corresponding optimal expected performance index is
\[E(J)(u^*(k)) = z^T(0)S(0)z(0) + \sum_{k=0}^{N-1} \text{Tr}[S(k+1)\Theta_k],\]
where \(\text{Tr}[]\) denotes the trace of matrix.

It is clear that the traditional LQG theory is to minimize the expected value of the performance index, \(J\), only concerning a sole objective optimization. Note that each realization of the control policy can only result in a single value of the performance index. Thus, the distribution of the realized values of the performance index could range widely. It is, in many situations, necessary to control both the expected value and the dispersion of the performance index for the LQG problem.

In general, the degree of the dispersion of the randomly valued \(J\) can be approximately measured by its variance \(\text{Var}\{J\}\), thus, we consider the following optimal control problem
\[(P) \quad \min_{u(k)} \phi(E(J), \text{Var}(J)) \quad \text{s.t.} \quad z(k+1) = A(k)z(k) + B(k)u(k) + \xi(k) \\
k = 0, 1, \ldots, N-1, z(0) = z_0,
\]
where the overall objective function, \(\phi\), is a continuous differentiable nonlinear function of the mean \(E(J)\) and the variance \(\text{Var}(J)\), and is assumed to be a strictly convex function, at the same time we also require that \(\phi\) is a strictly increasing function of \(E(J)\) and \(\text{Var}(J)\), i.e.,
\[\frac{\partial \phi}{\partial E(J)} > 0, \quad \frac{\partial \phi}{\partial \text{Var}(J)} > 0.\]

The interpretation of (5) is that improving each individual performance index leads to an improvement of the overall performance measure. In addition, the reason the \(\phi\) is assumed to be the convex function is explained as follows.

In investment science [18], an investor with a concave utility function is said to be risk-averse. This terminology is used because of the following, known as Jensen’s inequality, which states that if \(D\) is a concave function then, for any random variable \(X\),
\[E[D(X)] \leq D(E[X]).\]

Hence, letting \(X\) be the return from an investment, it follows from Jensen’s inequality that any investor with a concave utility function would prefer the certain return of \(E[X]\) to receiving a random return with this mean.

In control problem \((P)\), the overall objective function \(\phi\) reflects the decision maker’s utility function over the system. This is because the mean \(E(J)\) reflects the performance of the closed-loop system, and the variance term, \(\text{Var}(J)\), measures the degree of the dispersion of the randomly valued \(J\). Note that the utility function \(D\) for the investor of risk-averse is concave in the sense of the maximizing, thus, in control problem \((P)\), the utility function \(\phi\) for the controller of risk-averse is taken as convex function in the sense of the minimizing.

For writing convenience, we define \(J_E = E(J)\) and \(J_{\text{Var}} = \text{Var}(J)\). In this way, the overall objective function \(\phi\) in \((P)\) can be written as \(\phi(J_E, J_{\text{Var}})\). It is very difficult to directly solve problem \((P)\) because of the nonlinearity of \(\phi\) and the non-separability of the variance in the sense of dynamical programming.

The following auxiliary problem is now constructed for problem \((P)\) with a fixed parameters \(w \in R\)
\[(A(w)) \quad \min_{u(k)} J_E + wJ_{\text{Var}} \quad \text{s.t.} \quad z(k+1) = A(k)z(k) + B(k)u(k) + \xi(k) \\
k = 0, 1, \ldots, N-1, z(0) = z_0,
\]
where parameter \(w\) represents the trade-off between the expected value and the variance of the performance index \(J\). The larger the value of \(w\), the more the importance has been placed on the dispersion control.

Theorem 1: Suppose \(\{u^*(k)\}\) is an optimal control of the auxiliary problem \((A(w^*))\), then \(\{u^*(k)\}\) is also an optimal control of problem \((P)\), when \(w^*\) satisfies
\[w^* = \frac{\partial \phi^*(w^*)}{\partial J_E},\]
The implication of Theorem 1 is that any optimal solution to problem \((P)\) is in the set of solutions to auxiliary problem \((A(w^*))\).

III. SOLUTION OF AUXILIARY PROBLEM

The results in this section are identical to what appeared in [7], here, for convenient of reader we briefly list them. It is well know that dynamic programming is the only universal solution for stochastic control that generates a closed-loop policy. More specifically, when the performance index satisfies the separability and monotonicity, especially, when the performance index is defined as (4), then the
following powerful decomposition holds due to the smooth property of the expectation operator,
\[
\min_{u(k)} E[J(z, u)|I^0] =
\min_{u(k)} E\{E[\ldots E[E[J(z, u)|I^{N-1}]|I^{N-2}]\ldots |I^1]|I^0\}
\]
\[
= \min_{u(0)} J_0(z(0), u(0)) + E\{ \min_{u(1)} J_1(z(1), u(1))
+ E\{ \ldots E\min_{u(N-2)} \{ J_{N-1}(z(N-2), u(N-2))
+ E\{ \min_{u(N-1)} J_{N-1}(z(N-1), u(N-1))
+ EJ_N(z(N)) |I^{N-1}\} |I^{N-2}\} |I^1\} |I^0\},
\]
where \(I^k\) is the information set at stage \(k\), i.e.,
\[
I^k = \{z(0), u(0), z(1), u(1), \ldots, z(k), u(k)\}.
\]
The major difficulty in obtaining an optimal closed-loop policy in auxiliary problem is due to the non-separability of the variance, more specifically, due to the smooth property does not apply to the operator of variance:

\[\text{Var}(\cdot | I^t) \neq \text{Var}[\text{Var}(\cdot) | I^t], \forall t < k.\]  

The key issue to get a closed-loop optimal control policy for the mean-variance control of linear-quadratic Gaussian problems is how to get a nested recursive relationship which could be exploited by dynamic programming. Notice that the following decomposition formula holds for the variance, 

\[\text{Var}(\cdot | I^t) = \text{Var}[\text{Var}(\cdot | I^k) | I^t] + E[\text{Var}(\cdot | I^k) | I^t], \forall t < k.\]  

Define the cost-to-go at stage \(t\) as

\[
J_{t-N} = z^T(N)Q(N)z(N) + \sum_{k=t}^{N-1} [z^T(k)Q(k)z(k) + u^T(k)R(k)u(k)]
\]

and

\[
J_{N-N} = z^T(N)Q(N)z(N).
\]

Notice that for \(t \leq k\),

\[\text{Var}(J_{t-N} | I^k) = \text{Var}(J_{(k+1)-N} | I^k).\]

Then, from (8) the following nested form can be then derived for the variance of performance index \(J\)

\[\text{Var}(J | I^0) = E[\text{Var}(J_{2-N} | I^1) | I^0]
+ \text{Var}[E(\text{Var}(J_{3-N} | I^2) | I^1) | I^0)
+ E\{ \text{Var}[E(\text{Var}(J_{4-N} | I^3) | I^2) | I^1)| I^0)\}
+ \ldots
+ E\{ \text{Var}[E(\text{Var}(J_{N-N} | I^{N-1}) | I^{N-2}) | I^1)| I^0)\}
+ \sum_{t=1}^{N-1} E\{ \text{E}[E(\text{Var}(J_{t-N} | I^t) | I^{t-1}) | I^{t-2}) | I^1)| I^0)\}.\]  

The control policy in this paper is assumed of a linear feedback form

\[u(k) = K(k)z(k),\]  

where feedback gain matrix \(K(k)\) is to be determined. Using the mathematical induction, we can prove the following results,

\[E[J_{t-N} | I^t] = z^T(t)\Phi_tz(t) + \sum_{k=t+1}^{N} \text{Tr}[\Phi_k\Theta_k],\]

where \(\Phi_t\) satisfies the following backward recursion for \(k = N - 1, N - 2, \ldots, t,\)

\[\Phi_k = [A + BK(k)]^T \Phi_{k+1} [A + BK(k)]
+ K^T(k)R(k)K(k) + Q(k),\]

with the boundary condition \(\Phi_N = Q(N)\).

We can further get the mean and the variance of the quadratic term \(z^T(t + 1)\Phi_{t+1}z(t + 1)\) conditioned on \(I^t\) as follows,

\[E[z^T(t + 1)\Phi_{t+1}z(t + 1) | I^t] = z^T(t)[A + BK(t)]^T \Phi_{t+1}
\times [A + BK(t)]z(t) - \text{Tr}[\Phi_{t+1}\Theta]\]

and

\[\\text{Var}[z^T(t + 1)\Phi_{t+1}z(t + 1) | I^t] = E[\{z^T(t + 1)\Phi_{t+1}z(t + 1)
- E(z^T(t + 1)\Phi_{t+1}z(t + 1) | I^t)\}^2 | I^t]
= 4z^T(t)[A + BK(t)]^T \Phi_{t+1}\Theta_{t+1}
\times [A + BK(t)]z(t) - \text{Tr}[\Phi_{t+1}\Theta]\}
+ E\{\{z^T(t)\Phi_{t+1}\xi(t)\}^2\}.\]

Through some mathematical derivations. The variance in (13) can be now rewritten as,

\[\\text{Var}[z^T(t + 1)\Phi_{t+1}z(t + 1) | I^t] = E[\{z^T(t + 1)\Phi_{t+1}z(t + 1)
- E(z^T(t + 1)\Phi_{t+1}z(t + 1) | I^t)\}^2 | I^t]
= 4z^T(t)[A + BK(t)]^T \Phi_{t+1}\Theta_{t+1}
\times [A + BK(t)]z(t) + 2\text{Tr}[(\Phi_{t+1}\Theta)^2].\]

From (14), we have

\[\\text{Var}(J_{N-N} | I^{N-1}) = \text{Var}\{z^T(N)Q(N)z(N) | I^{N-1}\}
= z^T(N - 1)\Phi_{N-1}z(N - 1)
+ 2\text{Tr}[(\Phi_N\Theta)^2],\]

where

\[
\Phi_{N-1} = 4[A + BK(N - 1)]^T \Phi_N\Theta_N[4[A + BK(N - 1)].
\]
From (12) and (15), the first term in the decomposition formula (9) can be expressed as

\[
E\{E\{E[Var(J_{N-N} | I^{N-1}) | I^{N-2}] \ldots | I^{1} | I^{0}\}\}
\]

\[
= z^{T}(0) \Phi_{0}^{N-1} z(0) + \sum_{k=1}^{N} \text{Tr} [\Phi_{k}^{N-1} \Theta_{\xi}]
+ 2\text{Tr}(\Phi_{k}^{N} \Theta_{\xi}^{2}),
\]

(16)

where for \( k = N - 2, N - 3, \ldots, 0, \)

\[
\Phi_{k}^{N-1} = [A + BK(k)]^{T} \Phi_{k+1}^{N-1} [A + BK(k)]
\]

with the boundary condition \( \Phi_{N-1} = 0 \).

From (11), (12) and (14),

\[
Var[E\{E[Var(E(J_{N-N} | I^{t}) | I^{t-1}) | I^{t-2}] \ldots | I^{1} | I^{0}\}
\]

\[
= z^{T}(t-1) \Phi_{t-1}^{t-1} z(t-1) + 2\text{Tr}[(\Phi_{t} \Theta_{\xi})^{2}],
\]

(17)

Thus, using (17), the second term in the decomposition formula (9) can be expressed as the following,

\[
E\{E\{E[Var(E(J_{N-N} | I^{t}) | I^{t-1}) | I^{t-2}] \ldots | I^{1} | I^{0}\}
\]

\[
= z^{T}(0) \Phi_{0}^{t} z(0) + \sum_{k=1}^{t-1} \text{Tr} [\Phi_{k}^{t-1} \Theta_{\xi}]
+ 2\text{Tr}(\Phi_{k}^{t} \Theta_{\xi}^{2}),
\]

(18)

where \( \Phi_{k}^{t-1} \) satisfies the following backwards recursion,

\[
\Phi_{k}^{t-1} = [A + BK(k)]^{T} \Phi_{k+1}^{t-1} [A + BK(k)]
\]

\[ k = t - 2, t - 3, \ldots, 0 \]

with the boundary condition \( \Phi_{1-1} = 0 \).

From (11), we have

\[
E\{J | I^{0}\} = E\{J_{0-N} | I^{0}\}
\]

\[
= z^{T}(0) \Phi_{0} z(0) + \sum_{k=1}^{N} \text{Tr}[\Phi_{k} \Theta_{\xi}].
\]

(19)

Substituting (16) and (18) into (9), we have

\[
Var\{J\} = z^{T}(0) \sum_{t=1}^{N} \Phi_{0}^{t-1} z(0)
+ \sum_{t=2}^{N} \sum_{k=1}^{t-1} \text{Tr}[\Phi_{k}^{t-1} \Theta_{\xi}]
+ 2\sum_{t=1}^{N} \text{Tr}[(\Phi_{t} \Theta_{\xi})^{2}].
\]

(20)

Substituting (19) and (20) into \( E\{J\} + wVar\{J\} \), the objective function in auxiliary problem \( (A(w)) \), we have

\[
\hat{\phi} = E\{J\} + wVar\{J\}
\]

\[
= z^{T}(0) \Phi_{0} z(0) + \sum_{k=1}^{N} \text{Tr}[\Phi_{k} \Theta_{\xi}]
+ w\left\{ z^{T}(0) \sum_{t=1}^{N} \Phi_{0}^{t-1} z(0)
+ \sum_{t=2}^{N} \sum_{k=1}^{t-1} \text{Tr}[\Phi_{k}^{t-1} \Theta_{\xi}]
+ 2\sum_{t=1}^{N} \text{Tr}[(\Phi_{t} \Theta_{\xi})^{2}] \right\}.
\]

(21)

Recall that the initial state \( z(0) \) is exactly known. Since every term of \( \hat{\phi} \) in (21) is a function of the feedback gain matrices at different stages, \( K(t), t = 0, 1, \ldots, N - 1 \), the objective function in \( (A(w)) \) now becomes a function of \( K(t) \), namely,

\[
\hat{\phi} = \hat{\phi}[K(0), K(1), \ldots, K(N - 1)].
\]

(22)

Thus, the auxiliary problem \( (A(w)) \) with a given \( w \) is now reduced to the following deterministic unconstrained minimization problem with respect to \( K(0), K(1), \ldots, K(N-1) \):

\[
\text{(DMV)} \quad \min \hat{\phi}[K(0), K(1), \ldots, K(N - 1)]
\]

Notice that the above deterministic unconstrained minimization problem \( (DMV) \) does not depend on the state \( z(k) \) at various stages. Thus, \( (DMV) \) can be solved offline. Once we know the optimal gain matrices \( K(0), K(1), \ldots, K(N - 1) \) generated by solving \( (DMV) \), the optimal closed-loop control law given in (10) can be then determined.

IV. DETERMINATION OF THE OPTIMAL PARAMETER \( w \)

Due to the complexity of \( \hat{\phi}(K(0), K(1), \ldots, K(N - 1)) \), it seems almost impossible to solve problem \( (DMV) \) analytically. We can, however, always adopt some efficient numerical schemes to solve this nonlinear programming problem. The optimal solution of problem \( (DMV) \) \( K(0), K(1), \ldots, K(N - 1) \) are functions of parameter \( w \). The optimal weight coefficient \( w^{*} \) can be obtained by gradient method in this paper.

Let \( \nabla \phi \) denote the gradient vector of \( \phi \) on the noninferior \( \{E(J), Var(J)\} \), then

\[
\nabla \phi = \left[ \frac{\partial \phi}{\partial J_{E}}, \frac{\partial \phi}{\partial Var} \right]^{T}.
\]

(23)

Set \( \tau = [1, w]^{T} \), and construct the following direction vector

\[
V(\tau) = [V_{1}(\tau), V_{2}(\tau)]^{T}
\]

\[
= - \nabla \phi + \frac{\tau^{T} \nabla \phi}{\tau^{T} \tau} \eta.
\]

(24)

If we let \( V(\tau) = 0 \), then we have the following equations:

\[
- \frac{\delta \phi}{\delta J_{E}} + \frac{\tau^{T} \nabla \phi}{\tau^{T} \tau} \eta = 0
\]

(25)

\[
- \frac{\delta \phi}{\delta Var} + \frac{\tau^{T} \nabla \phi}{\tau^{T} \tau} \eta = 0.
\]

(26)
Combining (25) and (26) yields the optimality condition (7).
We can prove the vector \( V(\tau) \) is a decent direction of the objective function \( \phi \) in the performance index space \( \{J_E, J_{\text{Var}}\} \).

In order to make \( V(\tau) = 0 \), the following \( \epsilon \)-constraint problem is formulated to realize a feasible descent direction

\[
\begin{align*}
\text{(EOP)} & \quad \min_{u(k)} J_E \\
\text{s.t.} & \quad J_{\text{Var}}(u) \leq J_{\text{Var}}(\hat{u}(\tau^m)) + \alpha V_2(\tau^m) \\
& \quad z(k+1) = A(k)z(k) + B(k)u(k) + \xi(k) \\
& \quad k = 0, 1, \ldots, N-1, \quad z(0) = z_0,
\end{align*}
\]

where \( \alpha \) is a strictly positive step-size parameter which can be adjusted during the iteration to guarantee a decrement of the overall objective function, \( \tau^m \) is the weighting vector iteration \( m \), and \( \hat{u}(\tau^m) \) is the optimal solution of auxiliary problem \( (A(w^m)) \).

The dual problem of \( \text{(EOP)} \) is

\[
\begin{align*}
\text{(DOP)} & \quad H(\lambda) = \min_{u(k)} L(u, \lambda) \\
\text{s.t.} & \quad J_{\text{Var}}(u) = J_{\text{Var}}(\hat{u}(\tau^m)) - \alpha V_2(\tau^m)
\end{align*}
\]

and \( \lambda \) is the nonnegative Lagrangian factor. Since \([J_{\text{Var}}(u) + \alpha V_2(\tau^m)]\) is constant, \( \text{(DOP)} \) is equivalent to auxiliary problem \( (A(w)) \) when \( w = \lambda \).

It is clear that \( \lambda \) in \( \text{(DOP)} \) can be selected to serve the same role as the weighting vector \( w^{m+1} \) at iteration \( m+1 \) in \( (A(w)) \). In order to realize a decent direction at iteration \( m+1 \), the new values of \( \lambda \) can be updated as follows:

\[
\lambda^{m+1} = \lambda^{m} + \frac{\partial H(\lambda^{m})}{\partial \lambda} = \lambda^{m} - \alpha V_2(\tau^m).
\]

Since the \( w \) in auxiliary problem \( (A(w)) \) has the same role with \( \lambda \) in dual problem. Thus, the value of \( w \) can be updated as (29), namely,

\[
w^{m+1} = w^{m} - \alpha V_2(w^m),
\]

where \( \alpha \) is the step size parameter.

V. SOLUTION SCHEME AND EXAMPLES

The derived closed-loop control law for the optimal trade-off LQG problems can be now integrated with a search scheme for the optimal parameter \( w^* \) to control problem \( (P) \). Notice that both \( E(J) \) and \( \text{Var}(J) \) are \( w \) dependent. In order to check the optimality condition (7) for parameter \( w \), \( E(J) \) and \( \text{Var}(J) \) need to be calculated for given \( w \) by (19) and (20). An efficient numerical algorithm using the updating formula (30) can be devised such that parameter \( w \) converges to an optimal \( w^* \) that satisfies the optimal necessary condition (7) in an iterative process. The feedback optimal control law \( u^*(k) \) specified in equation (10) that corresponds to \( w^* \) solves the control problem \( (P) \). We now describe formally the solution algorithm for control problem \( (P) \).

Algorithm 1 Solution algorithm for problem \( (P) \):

**Step 0:** Choose a very small positive number \( \epsilon \) for error tolerance, set the iteration number \( m = 0 \) and give the initial value of \( w^m \);

**Step 1:** For given \( w^m \), using the gradient method to solve \( (DMV) \), obtain the gain matrix \( K(0), K(1), \ldots, K(N-1) \);

**Step 2:** Calculate the gradient vector \( \nabla \phi \) and the direction vector \( V(\tau) \) by (23) and (24);

**Step 3:** If the necessary optimality condition (7) is satisfied, namely, \( ||V(w^m)|| \leq \epsilon \), the solution \( u(k) \) to \( (DMV) \) is just the optimal control of \( (P) \), then stop. Otherwise, set \( m := m+1 \) and update \( w^m \) by equation (30) and go back to Step 1.

Algorithm 1 is independent of the state of system, thus, it is an off-line algorithm. After the optimal feedback control gains \( K^*(0), K^*(1), \ldots, K^*(N-1) \) are obtained, using the following algorithm can control system such that the mean and the variance of the performance index of the closed-loop system achieve the optimal trade-off in the sense of utility function \( \phi \).

Algorithm 2

**Step 0:** Set \( k = 0 \);

**Step 1:** Observe \( z(k) \), and calculate control \( u(k) = K^*(k)z(k) \);

**Step 2:** Apply \( u(k) \) to the system. If \( k = N-1 \), stop; Otherwise, set \( k := k+1 \) and go back to Step 1.

The following two stages scalar example illustrates the control results under different control framework.

**Example:** Consider the following two-stage scalar LQG problem

\[
\begin{align*}
\min & \quad \lambda^{m+1} = \lambda^{m} - \alpha V_2(\tau^m). \\
\text{s.t.} & \quad z(k+1) = A(k)z(k) + B(k)u(k) + \xi(k), \\
& \quad k = 0, 1, \ldots, N-1. \\
\end{align*}
\]

The performance index is

\[
J = \frac{z^T(N)Q(N)z(N)}{
+ \sum_{k=0}^{N-1} [z^T(k)Q(k)z(k) + u^T(k)R(k)u(k)],
\]

where \( A(k) = 0.8, B(k) = 0.5, z(0) = 0.5 , \xi(k) \sim N(0, \Theta_\xi) \) and \( \Theta_\xi = 0.5, Q = 1, R = 1 \). The first two stages is considered, namely \( N = 2 \).

A. Scheme 1: Classical LQG optimal controller

Under the classical LQG optimal control framework, the optimal control problem is the following

\[
\begin{align*}
\min & \quad E(J) \\
\text{s.t.} & \quad z(k+1) = A(k)z(k) + B(k)u(k) + \xi(k), \\
& \quad k = 0, 1. \\
\end{align*}
\]

The optimal control can be obtained by dynamic programming,

\[
u^*(k) = -\Gamma(k)z(k), \quad k = 0, 1,
\]
where
\[ \Gamma(0) = 0.4389, \quad \Gamma(1) = 0.32. \]
Simple computation gives the following values for the mean and the variance of performance index \( J \),
\[ E\{J\} = 1.6816 \]
\[ Var\{J\} = 2.5070. \]

B. Scheme 2: Mean-variance controller

Under the mean-variance control framework, the optimal control problem is the following
\[
\min E(J) + w Var(J)
\]
subject to
\[ z(k + 1) = A(k)z(k) + B(k)u(k) + \xi(k) \]
\[ k = 0, 1, \ldots, N - 1 \]
where \( w \) is a given weight coefficient in advance.

For given \( w \), using gradient method can search out the optimal solution of \( \phi \) with respect to \( K(0) \) and \( K(1) \).

C. Scheme 3: Optimal trade-off controller

The nonlinear utility function between the mean and the variance of performance index is taken as
\[ \phi = J_E^2 + \exp\{J_{Var}\}, \]
where \( J_E = E(J) \), \( J_{Var} = Var(J) \).

Set the initial value of \( w = 1 \), the step size in (30) \( \alpha = 0.2 \) and the error tolerance \( \epsilon = 0.05 \). Then, we use the gradient method to search the control gains \( K(0) \) and \( K(1) \). After twelve iterations \( \| V(w) \| \leq \epsilon \). This means that the objective function \( \phi(J_E, J_{Var}) \) is minimized, the optimal weighted coefficient \( w^* = 1.9989 \) for the auxiliary problem \( A(w) \), the optimal control gains are \( K^*(0) = -1.2061 \) and \( K^*(1) = -0.4473 \) and the error that is the absolute value of \( V^2(\tau^m) \) is 0.0432.

It is easily seen that the values of the mean of performance index \( J \) are 1.6816 and 1.8946 for the LQG controller and the optimal trade-off controller, respectively, and the values of variance are 2.507 and 2.0581, respectively. A 12.67 percent increase in the mean is traded for a 17.91 percent decrease in the variance over that of the LQG solution. The improvement of performance index depends on the selection of the utility function \( \phi \).

VI. Conclusion

In the traditional discrete-time LQG control problem, only the expected value of the performance index is under control. In applications, it is often the case that the dispersion of the performance index spans widely which is not desirable in many situations. In this paper, a kind of optimal trade-off performance index spans widely which is not desirable in the expected value of the performance index is under control. The nonlinear utility function between the mean and the variance of performance index depends on the selection of the utility function \( \phi \).

REFERENCES