The Geometry of Quadratic Programming

María M. Seron

September 2004
1. Introduction

2. The Geometry of Quadratic Programming
   - QP Solution for $N = 2$
   - QP Solution for Arbitrary $N$

3. Solution Using the KKT Optimality Conditions
In the previous lecture we saw how dynamic programming can be used to derive a simple parameterisation of a finite horizon optimal control problem and its RHC implementation.

Here we will analyse the geometric structure of the finite horizon optimal control problem when seen as a quadratic programme [QP].

Using this geometric structure, we will re-derive in the next lecture the result obtained via dynamic programming.

We start with a simple case of horizon $N = 2$ and input constraints and then show how the same ideas extend to more general cases.
The Finite Horizon Optimal Control Problem

Let the system be given by

\[ x_{k+1} = Ax_k + Bu_k, \quad |u_k| \leq \Delta, \]  

(1)

where \( x_k \in \mathbb{R}^n \) and \( \Delta > 0 \) is the input constraint level. Consider the following fixed horizon optimal control problem

\[ \mathcal{P}_2(x) : \quad V_2^{\text{OPT}}(x) \triangleq \min V_2(\{x_k\}, \{u_k\}), \]  

(2)

subject to:

\[ x_{k+1} = Ax_k + Bu_k \quad \text{for} \ k = 0, 1, \]

\[ x_0 = x, \]

\[ u_k \in \mathbb{U} \triangleq [-\Delta, \Delta] \quad \text{for} \ k = 0, 1, \]

where the objective function in (2) is

\[ V_2(\{x_k\}, \{u_k\}) \triangleq \frac{1}{2} x_2^T P x_2 + \frac{1}{2} \sum_{k=0}^{1} \left( x_k^T Q x_k + u_k^T R u_k \right). \]  

(3)
To compare with the results previously obtained using Dynamic Programming, in the objective function

\[ V_2(\{x_k\}, \{u_k\}) \triangleq \frac{1}{2} x_2^T P x_2 + \frac{1}{2} \sum_{k=0}^{1} \left( x_k^T Q x_k + u_k^T R u_k \right), \]

we select \( Q > 0, R > 0 \) and \( P \) satisfying the algebraic Riccati equation

\[ P = A^T PA + Q - K^T \bar{R} K, \tag{4} \]

where

\[ K \triangleq \bar{R}^{-1} B^T PA, \quad \bar{R} \triangleq R + B^T PB. \tag{5} \]
For the above problem, the corresponding QP optimal solution is

$$\text{QP: } \mathbf{u}^{\text{OPT}}(x) = \arg\min_{\mathbf{u} \in R_{\text{uc}}} \frac{1}{2} \mathbf{u}^T H \mathbf{u} + \mathbf{u}^T F x,$$  \hspace{1cm} (6)

where

$$\mathbf{u} = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}, \quad H = \bar{R} \begin{bmatrix} 1 + (KB)^2 & KB \\ KB & 1 \end{bmatrix}, \quad F = \bar{R} \begin{bmatrix} K + KBKA \\ KA \end{bmatrix},$$ \hspace{1cm} (7)

and $R_{\text{uc}}$ is the square

$$- \begin{bmatrix} \Delta \\ \Delta \end{bmatrix} \leq \mathbf{u} \leq \begin{bmatrix} \Delta \\ \Delta \end{bmatrix}.$$

Note that the Hessian $H$ is positive definite since $R > 0$ and hence $\bar{R} = R + B^T PB > 0$. 
The QP in (6) has a nice geometric interpretation in the $u$-space. Consider the equation

$$\frac{1}{2} u^T H u + u^T F x = c, \quad c \text{ constant}. \quad (8)$$

This defines *ellipsoids* in $\mathbb{R}^2$ centred at $u_{uc}^{opt}(x) = -H^{-1} F x$.

Then the QP in (6) amounts to finding the *smallest* ellipsoid that intersects the boundary of $R_{uc}$, and $u^{opt}(x)$ is the *point of intersection*.
Coordinate Transformation

The problem can be simplified if we make a coordinate transformation via the square root of the Hessian, that is,

$$ \tilde{u} = H^{1/2} u. $$

(9)

In the new $\tilde{u}$-coordinates (9), the constraint set $R_{uc}$ is mapped into another set (denoted also by $R_{uc}$).

The ellipsoids (8) are now circles centred at $\tilde{u}_{uc}^{\text{OPT}}(x) = -H^{-1/2}Fx$.

Thus the QP (6) is transformed into the problem of finding the point in $R_{uc}$ that is closest to $\tilde{u}_{uc}^{\text{OPT}}(x)$ in the Euclidean distance.

This is qualitatively illustrated in the following figure.
In the new coordinates, the level sets are circles in $\mathbb{R}^2$ centred at $\tilde{u}_{\text{opt}}(x) = -H^{-1/2}Fx$.

Then the QP in (6) amounts to finding the point in $R_{\text{uc}}$ that is closest to $\tilde{u}_{\text{opt}}(x)$ in the Euclidean distance.

This transformed geometric picture allows us to immediately write down the solution to the QP for this special case.
The solution of the QP is obtained by partitioning $\mathbb{R}^2$ into nine regions.

The first region is the polytope $R_{uc}$.

Regions $R_1$ to $R_8$ are delimited by lines that are normal to the faces of $R_{uc}$ and pass through its vertices.
The optimal constrained solution $\tilde{u}^{OPT}(x)$ is determined by the region in which the optimal unconstrained solution $\tilde{u}_{uc}^{OPT}(x)$ lies, in the following way:

- First, the solution in $R_{uc}$ is $\tilde{u}^{OPT}(x) = \tilde{u}^{OPT}_{uc}(x)$;
- Next, the solution in regions $R_1$, $R_3$, $R_5$ and $R_7$ is simply equal to the vertex that is contained in the region.
- Finally, the solution in regions $R_2$, $R_4$, $R_6$ and $R_8$ is defined by the orthogonal projection of $\tilde{u}^{OPT}_{uc}(x)$ onto the faces of $R_{uc}$. 

Centre for Complex Dynamic Systems and Control
As a result of above procedure, we obtain a characterisation of the QP solution in an active region $R_i$ as

$$u^{\text{opt}}(x) = H^{-1/2}\tilde{u}^{\text{opt}}(x) \quad \text{if} \quad \tilde{u}_{uc}^{\text{opt}}(x) \in R_i,$$

where $\tilde{u}^{\text{opt}}(x)$ is an affine function of $x$ and $\tilde{u}_{uc}^{\text{opt}}(x) = -H^{-1/2}Fx$.

We next proceed to find an algebraic description of a particular region and an expression for $u^{\text{opt}}(x)$ in it.
Consider the active region $R_8$ associated with face $f_4$.

It is delimited by face $f_4$ and its normals $n_1$ and $n_4$ passing through the vertices $v_1$ and $v_4$.

The line that contains face $f_4$ corresponds to the second control $u_1$ equal to the saturation level $\Delta$, that is $u_1 = [0 \ 1]u = \Delta$ in the original $u$-coordinates.

In the new $\tilde{u}$-coordinates (9), it is then defined by the equation

$$f_4 : [0 \ 1]H^{-1/2}\tilde{u} = \Delta.$$
Hence, face $f_4$ is:

$$f_4 : \begin{bmatrix} 0 & 1 \end{bmatrix} H^{-1/2} \tilde{u} = \Delta.$$ 

The normal lines $n_1$ and $n_2$ are:

$$n_1 : \begin{bmatrix} 1 & 0 \end{bmatrix} H^{1/2} \tilde{u} = H_1 \Delta \begin{bmatrix} 1 & 1 \end{bmatrix}^T.$$ 

$$n_4 : \begin{bmatrix} 1 & 0 \end{bmatrix} H^{1/2} \tilde{u} = H_1 \Delta \begin{bmatrix} -1 & 1 \end{bmatrix}^T.$$ 

Combining the above three equations, region $R_8$ is

$$R_8 : \begin{cases} [0, 1] H^{-1/2} \tilde{u} \geq \Delta \\
H_1 \Delta \begin{bmatrix} -1 & 1 \end{bmatrix}^T \leq [1, 0] H^{1/2} \tilde{u} \leq H_1 \Delta \begin{bmatrix} 1 & 1 \end{bmatrix}^T \end{cases}$$
The optimal constrained solution in $R_8$ is given by the normal projection of the unconstrained solution $\tilde{\mathbf{u}}_{\text{opt}}(x)$ onto face $f_4$.

That is, $\tilde{\mathbf{u}}_{\text{opt}}(x)$ satisfies the equation of face $f_4$ and that of the normal to it passing through $\tilde{\mathbf{u}}_{\text{opt}}(x)$:

$$[0 \ 1]H^{-1/2}\tilde{\mathbf{u}}_{\text{opt}}(x) = \Delta,$$

$$[1 \ 0]H^{1/2}\tilde{\mathbf{u}}_{\text{opt}}(x) = [1 \ 0]H^{1/2}\tilde{\mathbf{u}}_{\text{opt}}(x).$$
Hence, $\tilde{u}^{\text{OPT}}(x)$ satisfies:

$$
\begin{bmatrix}
0 & 1
\end{bmatrix} H^{-1/2} \tilde{u}^{\text{OPT}}(x) = \Delta,
$$

$$
\begin{bmatrix}
1 & 0
\end{bmatrix} H^{1/2} \tilde{u}^{\text{OPT}}(x) = \begin{bmatrix}
1 & 0
\end{bmatrix} H^{1/2} \tilde{u}_{\text{UC}}^{\text{OPT}}(x).
$$

Using $\tilde{u}^{\text{OPT}} = H^{1/2} u^{\text{OPT}} = H^{1/2} [u_0^{\text{OPT}}(x) \quad u_1^{\text{OPT}}(x)]^T$ and $\tilde{u}_{\text{UC}}^{\text{OPT}}(x) = -H^{-1/2} F x$ in the above equations, and solving for the first component $u_0^{\text{OPT}}(x)$ yields

$$
u_0^{\text{OPT}}(x) = -G x - h. \tag{10}$$

where

$$G = \frac{K + KBKA}{1 + (KB)^2} \quad \quad h = \frac{KB}{1 + (KB)^2} \Delta.$$  

Note that the above matrices coincide with the previously derived using Dynamic Programming.
Recapitulation

The solution in active region $R_8$ is

$$\begin{equation}
\mathbf{u}_{\text{opt}}(x) = \begin{bmatrix}
-Gx - h \\
\Delta
\end{bmatrix}
\end{equation}$$

if

$$R_8 : \begin{cases}
[0 \ 1]\mathbf{u}_{\text{uc}}^{\text{opt}}(x) \geq \Delta \\
H_1 \Delta [-1 \ 1]^T \leq [1 \ 0]H\mathbf{u}_{\text{uc}}^{\text{opt}}(x) \leq H_1 \Delta [1 \ 1]^T.
\end{cases}$$

where $\mathbf{u}_{\text{uc}}^{\text{opt}}(x) = -H^{-1}Fx$.

A similar characterisation can be derived in the remaining regions.
We will now consider arbitrary horizon $N$, $m \geq 1$ inputs, and more general constraints, leading to a QP of the form

$$u^{\text{opt}}(x) = \arg \min_{\Phi u \leq \Delta - \Lambda x} \frac{1}{2} u^T H u + u^T F x,$$  \hspace{1cm} (11)

where $u \in \mathbb{R}^{Nm}$, and where the constraint set

$$\Phi u \leq \Delta - \Lambda x, \quad \Phi \in \mathbb{R}^{q \times Nm}$$  \hspace{1cm} (12)

is a polyhedron in $\mathbb{R}^{Nm}$.
Again, it is convenient to work in the $\tilde{u}$-coordinates $\tilde{u} = H^{1/2}u$ to turn the QP into a minimum Euclidean distance problem.

Hence the constraint set, $R_{uc}$, is the polyhedron given by:

$$\tilde{\Phi} \tilde{u} \leq \Delta - \Lambda x, \quad \tilde{\Phi} \triangleq \Phi H^{-1/2}.$$
Active Regions

The procedure is the same as for $N = 2$, that is, we associate an *active region* to each face of the constraint set (before we associated region $R_8$ to face $f_4$). The active region is determined by the constraints that are active on the face. For example:

Face $f_1$ (one active constraint) and associated active region $R_1$.

Edge $e_1$ (two active constraints) and associated active region $R_2$. 

[Diagram showing active regions $R_1$ and $R_2$.]
Recall the constraint set

\[ \tilde{\Phi} \tilde{u} \leq \Delta - \Lambda x, \quad \tilde{\Phi} \in \mathbb{R}^{q \times Nm}. \]

A face is determined by a subset of the above \( q \) inequality constraints that are active, that is, they hold as equality constraints:

\[ \tilde{\Phi}_\ell \tilde{u} = \Delta_\ell - \Lambda_\ell x, \quad (13) \]

\( \ell \subset \{1, 2, \ldots, q\} \) is the set of indices of active constraints and the notation \( M_\ell \) indicates selection of rows of \( M \) with indices in \( \ell \).

The active region associated with face (13) is

\[
\begin{aligned}
&\begin{bmatrix}
\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T
\end{bmatrix}^{-1} \tilde{\Phi}_\ell \tilde{u} \leq 
\begin{bmatrix}
\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T
\end{bmatrix}^{-1} \begin{bmatrix}
\Delta_\ell - \Lambda_\ell x
\end{bmatrix},
\end{aligned}
\]

\[
\begin{aligned}
&\begin{bmatrix}
\tilde{\Phi}_s[1 - \tilde{\Phi}_\ell^T\tilde{\Phi}_\ell\tilde{\Phi}_\ell^T]^{-1} \tilde{\Phi}_\ell
\end{bmatrix} \tilde{u} \leq 
\begin{bmatrix}
\Delta_s - \Lambda_s x
\end{bmatrix} - \begin{bmatrix}
\tilde{\Phi}_s \tilde{\Phi}_\ell^T
\end{bmatrix} \begin{bmatrix}
\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T
\end{bmatrix}^{-1} \begin{bmatrix}
\Lambda_\ell x - \Delta_\ell
\end{bmatrix},
\end{aligned}
\quad (14)
\]

where \( s \) is the set of indices of inactive constraints.
The optimal constrained solution $\tilde{u}_{\text{opt}}(\bar{x})$ in the above active region is given by the point of the corresponding face that is closest to the unconstrained solution $\tilde{u}_{\text{uc}}(\bar{x}) = -H^{-1/2}Fx$.

This point is given by

$$\tilde{u}_{\text{opt}}(\bar{x}) = \tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}(\Delta \ell - \Lambda \ell \bar{x}) - [I - \tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}\tilde{\Phi}_\ell]H^{-1/2}Fx.$$  

In the original coordinates $u_{\text{opt}} = H^{-1/2}\tilde{u}_{\text{opt}}$ we have

$$u_{\text{opt}}(\bar{x}) = H^{-1/2}\tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}(\Delta \ell - \Lambda \ell \bar{x})$$

$$- H^{-1/2}[I - \tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}\tilde{\Phi}_\ell]H^{-1/2}Fx,$$

$$\triangleq G_\ell \bar{x} + h_\ell.$$
The solution in an active region is

\[
\tilde{u}^{\text{OPT}}(x) = \Phi_{\ell}^T[\Phi_{\ell} \tilde{\Phi}_{\ell}^T]^{-1}(\Delta_{\ell} - \Lambda_{\ell} x) - [I - \Phi_{\ell}^T[\Phi_{\ell} \tilde{\Phi}_{\ell}^T]^{-1} \Phi_{\ell}]H^{-1/2}Fx
\]

\[
\triangleq H^{1/2}(G_{\ell}x + h_{\ell}).
\]

if

\[
\begin{cases}
[\Phi_{\ell} \tilde{\Phi}_{\ell}^T]^{-1} \Phi_{\ell} \tilde{u}_{\text{OPT}}^{\text{OPT}}(x) & \leq [\Phi_{\ell} \tilde{\Phi}_{\ell}^T]^{-1}[\Delta_{\ell} - \Lambda_{\ell} x] \\
\tilde{\Phi}_{s}[I - \Phi_{\ell}^T[\Phi_{\ell} \tilde{\Phi}_{\ell}^T]^{-1} \Phi_{\ell}] \tilde{u}_{\text{OPT}}^{\text{OPT}}(x) & \leq -[\Lambda_{s} + \tilde{\Phi}_{s} \tilde{\Phi}_{\ell}^T[\tilde{\Phi}_{\ell} \tilde{\Phi}_{\ell}^T]^{-1} \Lambda_{\ell}]x + \Delta_{s} + \tilde{\Phi}_{s} \tilde{\Phi}_{\ell}^T[\tilde{\Phi}_{\ell} \tilde{\Phi}_{\ell}^T]^{-1} \Delta_{\ell},
\end{cases}
\]

where \(\tilde{u}_{\text{UC}}^{\text{OPT}}(x) = -H^{-1/2}Fx\) and \(\ell\) and \(s\) are the indices of active and inactive constraints, respectively.
The geometric structure of QP has allowed us to derive its solution in an active region.

The same solution can be derived using the KKT optimality conditions, as we show next.
The KKT Optimality Conditions

Consider the following nonlinear programming problem:

\[
\begin{align*}
\text{minimise} & \quad f(z), \\
\text{subject to:} & \quad g(z) \leq 0,
\end{align*}
\]

where \( f : \mathbb{R}^r \rightarrow \mathbb{R}, g = [g_1, \ldots, g_q]^T : \mathbb{R}^r \rightarrow \mathbb{R}^q \) are differentiable.

Let \( \bar{z} \) be a feasible solution, and let \( \ell = \{i : g_i(\bar{z}) = 0\} \).

Suppose that \( \nabla g_i(\bar{z})^T \) for \( i \in \ell \) are linearly independent and let \( \nabla g^T = [\nabla g_1^T, \ldots, \nabla g_q^T] \).

If \( \bar{z} \) is a local optimal solution, then there exist Lagrange multipliers \( \mu = [\mu_1, \ldots, \mu_q]^T \) such that the following KKT conditions hold:

\[
\nabla f(\bar{z})^T + \nabla g(\bar{z})^T \mu = 0, \\
\mu^T g(\bar{z}) = 0, \\
\mu \geq 0.
\]

(KKT)
KKT Conditions for QP

For the QP problem

\[
\text{minimise } \frac{1}{2} z^T H z + z^T c, \\
\text{subject to: } L z \leq W,
\]

the KKT conditions (KKT) are:

- **Primal Feasibility (PF):** \( L \bar{z} \leq W, \)
- **Dual Feasibility (DF):** \( H \bar{z} + c + L^T \mu = 0, \)
  \( \mu \geq 0, \)
- **Complementary Slackness (CS):** \( \mu^T (L \bar{z} - W) = 0. \)

(17)

If \( H \) is positive definite, the KKT conditions are necessary and sufficient for \( \bar{z} \) to be the unique global minimum.
Recall the QP (11) used in MPC:

In the original coordinates

\[
\begin{align*}
\text{minimise } & \quad \frac{1}{2} u^T H u + u^T F x, \\
\text{subject to: } & \quad \Phi u \leq \Delta - \Lambda x.
\end{align*}
\]

In the \(\tilde{u}\)-coordinates \(\tilde{u} = H^{1/2} u\)

\[
\begin{align*}
\text{minimise } & \quad \frac{1}{2} \tilde{u}^T \tilde{u} + \tilde{u}^T H^{-1/2} F x, \\
\text{subject to: } & \quad \tilde{\Phi} \tilde{u} \leq \Delta - \Lambda x.
\end{align*}
\]

When solved for different values of \(x\), the above problem is referred to as a multiparametric quadratic program [mp-QP].

An mp-QP is a QP in which the linear term in the objective function and the right hand side of the constraints depend linearly on a vector of parameters (the state vector \(x\) in this case).
Let $\tilde{u}^{\text{opt}}(x)$ be the optimal solution of the above problem. The KKT conditions (17) then are

\begin{align*}
\tilde{\Phi}\tilde{u}^{\text{opt}}(x) + \Lambda x &\leq \Delta, \quad \text{(PF)} \\
\tilde{u}^{\text{opt}}(x) + H^{-1/2}Fx + \tilde{\Phi}^T\mu &= 0, \quad \text{(DF1)} \\
\mu &\geq 0, \quad \text{(DF2)} \\
\mu^T[\tilde{\Phi}\tilde{u}^{\text{opt}}(x) - \Delta + \Lambda x] &= 0. \quad \text{(CS)}
\end{align*}

Consider now an active set of indices $\ell$ for $\tilde{u}^{\text{opt}}(x)$, and let $s$ be the corresponding inactive set. We then have that $\tilde{u}^{\text{opt}}(x)$ satisfies the equality constraints (13), that is,

\[ \tilde{\Phi}_\ell\tilde{u}^{\text{opt}}(x) = \Delta_\ell - \Lambda_\ell x. \]  

(18)

From (18), (CS) and (DF2), we have that

\[ \mu_s = 0, \quad \mu_\ell \geq 0. \]
Using $\mu_s = 0, \mu_\ell \geq 0$ in the (first) dual feasibility condition (DF1) and solving for $\tilde{u}^{\text{OPT}}(x)$ yields

$$\tilde{u}^{\text{OPT}}(x) = -H^{-1/2}Fx - \tilde{\Phi}_\ell^T\mu_\ell. \quad (19)$$

Using (19) in the active constraint equality (18) and solving for $\mu_\ell$ gives

$$\mu_\ell = [\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}[-\tilde{\Phi}_\ell H^{-1/2}Fx + \Lambda_\ell x - \Delta_\ell]. \quad (20)$$

Substituting the above equation into (19) we obtain

$$\tilde{u}^{\text{OPT}}(x) = \tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}(\Delta_\ell - \Lambda_\ell x) - [I - \tilde{\Phi}_\ell^T[\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1}\tilde{\Phi}_\ell]H^{-1/2}Fx, \quad (21)$$

which is identical to (15).
The inequalities (16) that define the region in the state space where the solution (21) is valid can be recovered as follows:

Combining the expression for the Lagrange multipliers (20)

\[
\mu_\ell = [\tilde{\Phi}_\ell \tilde{\Phi}_\ell^T]^{-1} [-\tilde{\Phi}_\ell H^{-1/2} Fx + \Lambda_\ell x - \Delta_\ell].
\]

and the sign condition \( \mu_\ell \geq 0 \) yields the first inequality in (16).

Finally, the second inequality in (16) follows from primal feasibility (see (PF)) of the inactive constraints, that is,

\[
\tilde{\Phi}_s \tilde{\mu}^{\text{OPT}}(x) + \Lambda_s x \leq \Delta_s,
\]

upon substitution of the expression (21) for the optimal solution.
In summary, we can see that the geometric characterisation given before is a particular arrangement of the KKT optimality conditions.

In particular:

- The halfspace above a face (for example, $f_4$) is given by the nonnegativity of the Lagrange multipliers corresponding to active constraints on the face.
- The band between normals to the face (for example, between $n_1$ and $n_4$) is given by feasibility of the inactive constraints on the face.