Lecture Notes: (Stochastic) Optimal Control

Marc Toussaint
Machine Learning & Robotics group, TU Berlin
Franklinstr. 28/29, FR 6-9, 10587 Berlin, Germany

July 1, 2010

Disclaimer: These notes are not meant to be a complete or comprehensive survey on Stochastic Optimal Control. This is more of a personal script which I use to keep an overview over control methods and their derivations. One point of these notes is to fix a consistent notation and provide a coherent overview for these specific methods.

Contents

1 Notation

x	system state (can be q or $(q, \dot q)$)
$q \in \mathbb{R}^n$	robot posture (vector of joint angles)
$y \in \mathbb{R}^d$	a task variable (e.g., endeffector position)
$\phi :\ q \mapsto y$	differentiable kinematic function
$J(q) = \frac{\partial \phi}{\partial q} \in \mathbb{R}^{d\times n}$	task Jacobian in posture q

We define a Gaussian over x with mean a and covariance matrix A as the function

$$\mathcal{N}(x \,|\, a, A) = \frac{1}{|2\pi A|^{1/2}}\ \exp\{-\frac{1}{2}(x-a)^\top A^{\text{-}1}\ (x-a)\} \tag{1}$$

with property $N(x \,|\, a, A) = N(a |\ x, A)$. We also define the canonical representation

$$\mathcal{N}[x \,|\, a, A] = \frac{\exp\{-\frac{1}{2}a^\top A^{\text{-}1} a\}}{|2\pi A^{\text{-}1}|^{1/2}}\ \exp\{-\frac{1}{2}x^\top A\, x + x^\top a\} \tag{2}$$

with properties

$$\mathcal{N}[x \,|\, a, A] = \mathcal{N}(x \,|\, A^{\text{-}1} a, A^{\text{-}1})\ ,\ \mathcal{N}(x \,|\, a, A) = \mathcal{N}[x \,|\, A^{\text{-}1} a, A^{\text{-}1}]\ .$$

The product of two Gaussians can be expressed as

$$\mathcal{N}[x\,|\,a,A]\;\mathcal{N}[x\,|\,b,B] = \mathcal{N}[x\,|\,a+b, A+B]\;\mathcal{N}(A^{\text{-}1}a\,|\,B^{\text{-}1}b, A^{\text{-}1}+B^{\text{-}1})\;, \tag{3}$$

$$\mathcal{N}(x\,|\,a,A)\;\mathcal{N}(x\,|\,b,B) = \mathcal{N}[x\,|\,A^{\text{-}1}a+B^{\text{-}1}b, A^{\text{-}1}+B^{\text{-}1}]\;\mathcal{N}(a\,|\,b, A+B)\;, \tag{4}$$

$$\mathcal{N}(x\,|\,a,A)\;\mathcal{N}[x\,|\,b,B] = \mathcal{N}[x\,|\,A^{\text{-}1}a+b, A^{\text{-}1}+B]\;\mathcal{N}(a\,|\,B^{\text{-}1}b, A+B^{\text{-}1})\;. \tag{5}$$

Linear transformations in x imply the following identities,

$$\mathcal{N}(Fx+f\,|\,a,A) = \frac{1}{|F|}\;\mathcal{N}(x\,|\;F^{\text{-}1}(a-f),\;F^{\text{-}1}AF^{\text{-}\top})\;, \tag{6}$$

$$= \frac{1}{|F|}\;\mathcal{N}[x\,|\;F^\top A^{\text{-}1}(a-f),\;F^\top A^{\text{-}1}F]\;, \tag{7}$$

$$\mathcal{N}[Fx+f\,|\,a,A] = \frac{1}{|F|}\;\mathcal{N}[x\,|\;F^\top(a-Af),\;F^\top AF]\;. \tag{8}$$

The joint Gaussian of two linearly dependent Gaussian variables reads

$$\mathcal{N}(x\,|\,a,A)\;\mathcal{N}(y\,|\,b+Fx,B) = \mathcal{N}\Big(\begin{pmatrix}x\\y\end{pmatrix}\Big|\begin{pmatrix}a\\b+Fa\end{pmatrix},\;\begin{pmatrix}A & A^\top F^\top\\ FA & B+FA^\top F^\top\end{pmatrix}\Big) \tag{9}$$

(See the lecture notes on "Gaussian identities" for more identities.)

Let us collect some matrix identities which we will need throughout. The *Woodbury identity*

$$(J^\top C^{\text{-}1}J+W)^{\text{-}1}J^\top C^{\text{-}1} = W^{\text{-}1}J^\top(JW^{\text{-}1}J^\top+C)^{\text{-}1}\;,$$

holds for any positive definite C and W. Further we have the identity

$$\mathbf{I}_n - (J^\top C^{\text{-}1}J+W)^{\text{-}1}J^\top C^{\text{-}1}J = (J^\top C^{\text{-}1}J+W)^{\text{-}1}W\;. \tag{10}$$

We define the pseudo-inverse of J w.r.t. W as

$$J^\sharp_W = W^{\text{-}1}J^\top(JW^{\text{-}1}J^\top)^{\text{-}1} \tag{11}$$

and similar quantity as

$$J^\natural_C = (JC^{\text{-}1}J^\top)^{\text{-}1}J^\top C^{\text{-}1}\;. \tag{12}$$

2 Stochastic optimal control (discrete time)

We assume a framework that is basically the same as for Markov Decision Processes but with a slight change in notation. Instead of maximizing rewards we will minimize costs; instead of an action a_t we refer to the control u_t; instead of $V^*(x)$ we denote the optimal value function by $J_t(x)$. Cost and dynamics are generally non-stationary.

Consider a discrete time stochastic controlled system

$$x_{t+1} = f_t(x_t,u_t)+\xi_t\;,\quad \xi_t\sim\mathcal{N}(0,Q_t) \tag{13}$$

with the state $x_t\in\mathbb{R}^n$, the control signal $u_t\in\mathbb{R}^m$, and Gaussian noise ξ of covariance Q. An alternative notation is

$$P(x_{t+1}\,|\,u_t,x_t) = \mathcal{N}(x_{t+1}\,|\,f_t(x_t,u_t),Q_t) \tag{14}$$

For a given state-control sequence $x_{0:T},u_{0:T}$ we define the cost

$$C(x_{0:T},u_{0:T}) = \sum_{t=0}^{T} c_t(x_t,u_t)\;. \tag{15}$$

Unlike the reward function in stationary MDPs, this cost function is typically not stationary. For instance, the cost might focus on the final state x_T relative to a goal state. In consequence also optimal policies are non-stationary. Just as we had in the MDP case, the value function obeys the Bellman optimality equation

$$J_t(x) = \min_u\Big[c_t(x,u)+\int_{x'}P(x'\,|\,u,x)\;J_{t+1}(x')\Big] \tag{16}$$

There are two versions of stochastic optimal control problems: The open-loop control problem is to "find a control sequence $u^*_{1:T}$ that minimizes the expected cost". The closed-loop (feedback) control problem is "find a control policy $\pi^*_t : x_t\mapsto u_t$ (that exploits the true state observation in each time step and maps it to a feedback control signal) that minimizes the expected cost".

2.1 The Linear-quadratic-Gaussian (LQG) case

Consider a **linear** control process with **Gaussian** noise,

$$P(x_{t+1} \mid x_t, u_t) = \mathcal{N}(x_{t+1} \mid A_t x_t + a_t + B_t u_t, Q_t) \ , \tag{17}$$

and **quadratic** costs,

$$c_t(x_t, u_t) = x_t^\top R_t x_t - 2 r_t^\top x_t + u_t^\top H_t u_t \ . \tag{18}$$

The general LQG case is specified by matrices and vectors $A_{0:T}, a_{0:T}, B_{0:T}, Q_{0:T}, R_{0:T}, r_{0:T}, H_{0:T}$. With a proper choice of R_t and r_t this corresponds to the problem of tracing an arbitrary desired trajectory x_t^*, where the cost is quadratic in $(x_t^* - x_t)$.

The LQG case allows us to derive an exact backward recursion, called Riccati equation, for the computation of the value function. The value function will always be a quadratic form of the state.

Let us assume that we know the value function $J_{t+1}(x)$ at time t+1 and that it has the form

$$J_{t+1}(x) = x^\top V_{t+1} x - 2 v_{t+1}^\top x \ . \tag{19}$$

Then

$$J_t(x) = \min_u \Big[x^\top R_t x - 2 r_t^\top x + u^\top H_t u + \int_y \mathcal{N}(y \mid A_t x + a_t + B_t u, Q_t)\ (y^\top V_{t+1} y - 2 v_{t+1}^\top y)\ dx \Big] \tag{20}$$

CONVENTION: For the remainder of this section we will drop the subscript $_t$ for A, a, B, Q, R, r, H – whereever it is missing we refer to time t.

The expectation of a quadratic form under a Gaussian is $\mathrm{E}_{\mathcal{N}(y \mid a, A)}\{y^\top V y - 2 v^\top y\} = a^\top V a - 2 v^\top a + \mathrm{tr}(VA)$. So we have

$$\begin{aligned} J_t(x) &= \min_u \Big[x^\top R x - 2 r^\top x + u^\top H u + (Ax + a + Bu)^\top V_{t+1} (Ax + a + Bu) \\ &\qquad - 2 v_{t+1}^\top (Ax + a + Bu) + \mathrm{tr}(V_{t+1} Q) \Big] \qquad (21) \\ &= \min_u \Big[x^\top (R + A^\top V_{t+1} A) x - 2 (r^\top + (v_{t+1} - V_{t+1} a)^\top A) x + u^\top (H + B^\top V_{t+1} B) u \\ &\qquad + 2 u^\top B^\top (V_{t+1} (Ax + a) - v_{t+1}) + a^\top V_{t+1} a - 2 v_{t+1}^\top a + \mathrm{tr}(V_{t+1} Q) \Big] \qquad (22) \end{aligned}$$

Minimizing w.r.t. u by setting the gradient to zero we have

$$\begin{aligned} 0 &= 2(H + B^\top V_{t+1} B) u^* + 2 B^\top (V_{t+1}(Ax + a) - v_{t+1}) \qquad (23) \\ u_t^*(x) &= -(H + B^\top V_{t+1} B)^{-1} B^\top (V_{t+1}(Ax + a) - v_{t+1}) \qquad (24) \\ J_t(x) &= x^\top (R + A^\top V_{t+1} A) x - 2 (r^\top + (v_{t+1} - V_{t+1} a)^\top A) x \\ &\quad + (V_{t+1}(Ax + a) - v_{t+1})^\top B (H + B^\top V_{t+1} B)^{-1} B^\top (V_{t+1}(Ax + a) - v_{t+1}) \\ &\quad - 2 (V_{t+1}(Ax + a) - v_{t+1})^\top B (H + B^\top V_{t+1} B)^{-1} B^\top (V_{t+1}(Ax + a) - v_{t+1}) \\ &\quad + a^\top V_{t+1} a - 2 v_{t+1}^\top a + \mathrm{tr}(V_{t+1} Q) \qquad (25) \\ J_t(x) &= x^\top V_t x - 2 x^\top v_t + \textit{terms independent of } x \ , \\ V_t &= R + A^\top V_{t+1} A - K V_{t+1} A \qquad (26) \\ v_t &= r + A^\top (v_{t+1} - V_{t+1} a) - K (v_{t+1} - V_{t+1} a) \qquad (27) \\ K &:= A^\top V_{t+1}^\top (V_{t+1} + (B H^{-1} B^\top)^{-1})^{-1} \ , \end{aligned}$$

The last equation for J_t is called the Ricatti equation. Initialized with $V_T = R_T$ and $v_T = r_T$ this gives a backward recursion to compute the value function J_t at each time step. Equation (24) also gives the optimal control policy.

Note that the optimal control and path is *independent of the process noise* Q.

2.2 Message passing in LQG

We may translate costs to probabilities by introducing a binary random variable z_t dependent on x_t and u_t,

$$P(z_t = 1 \mid u_t, x_t) = \exp\{-c_t(x_t, u_t)\} \ . \tag{28}$$

In the LQG case we can simplify this to

$$P(z_t\!=\!1 \,|\, x_t) \propto \mathcal{N}[x_t \,|\, r_t, R_t] \;, \tag{29}$$
$$P(u_t) = \mathcal{N}(u_t \,|\, 0, H^{\text{-}1}) \tag{30}$$

What is the posterior over the state trajectory $x_{1:T}$ conditioned on that we permanently observe $\hat c_t = 1$? Since we can integrate out the control u_t this is a simple Markov process with continuous state and Gaussian observation,

$$P(x_{t+1} \,|\, x_t) = \textstyle\int_u \text{d}u\; \mathcal{N}(x_{t+1} \,|\, Ax_t + a + Bu, Q)\; \mathcal{N}(u \,|\, 0, H^{\text{-}1}) \tag{31}$$
$$= \mathcal{N}(x_{t+1} \,|\, Ax_t + a, Q + BH^{\text{-}1}B^\top) \;. \tag{32}$$

Inference is a standard forward-backward process just as Kalman smoothing. The messages read

$$\begin{aligned}
\mu_{x_{t\text{-}1}\to x_t}(x_t) &= \mathcal{N}(x_t \,|\, s_t, S_t) \;,\\
s_t &= a_{t\text{-}1} + A_{t\text{-}1}(S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1}(S_{t\text{-}1}^{\text{-}1}s_{t\text{-}1} + r_{t\text{-}1})\\
S_t &= Q + B_{t\text{-}1}H^{\text{-}1}B_{t\text{-}1}^\top + A_{t\text{-}1}(S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1}A_{t\text{-}1}^\top
\end{aligned} \tag{33}$$
$$\begin{aligned}
\mu_{x_{t+1}\to x_t}(x_t) &= \mathcal{N}(x_t \,|\, v_t, V_t) \;,\\
v_t &= -A_t^{\text{-}1}a_t + A_t^{\text{-}1}(V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}(V_{t+1}^{\text{-}1}v_{t+1} + r_{t+1})\\
V_t &= A_t^{\text{-}1}[Q + B_tH^{\text{-}1}B_t^\top + (V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}]A_t^{\text{-}\top}
\end{aligned} \tag{34}$$
$$\mu_{\hat c_t\to x_t}(x_t) = \mathcal{N}[x_t \,|\, r_t, R_t] \;, \tag{35}$$

The potentials (v_t, V_T) which define the backward message can also be expressed in a different way: let us define

$$\bar V_{t+1} = V_{t+1}^{\text{-}1} + R_{t+1} \tag{36}$$
$$\bar v_{t+1} = V_{t+1}^{\text{-}1}v_{t+1} + r_{t+1} \;, \tag{37}$$

which corresponds to a backward message (in canonical representation) which has the cost message already absorbed. Using a special case of the Woodbury identity,

$$(A^{\text{-}1} + B)^{\text{-}1} = A - A(A + B^{\text{-}1})^{\text{-}1}A \;, \tag{38}$$

the bwd messages can be rewritten as

$$\begin{aligned}
V_{t+1}^{\text{-}1} &= A^\top[\bar V_{t+1}^{\text{-}1} + Q + BH^{\text{-}1}B^\top]^{\text{-}1}A\\
&= A^\top\bar V_{t+1}A - K\bar V_{t+1}A
\end{aligned} \tag{39}$$
$$K := A^\top\bar V_{t+1}[\bar V_{t+1} + (Q + BH^{\text{-}1}B^\top)^{\text{-}1}]^{\text{-}1}$$
$$\begin{aligned}
V_t^{\text{-}1}v_t &= -A^\top\bar V_{t+1}a_t + A^\top\bar v_{t+1} + K\bar V_{t+1}a_t - K\bar v_{t+1}\\
&= A^\top(\bar v_{t+1} - \bar V_{t+1}a_t) - K(\bar v_{t+1} - \bar V_{t+1}a_t)
\end{aligned} \tag{40}$$
$$\bar V_t = R_t + (A^\top - K)\bar V_{t+1}A \tag{41}$$
$$\bar v_t = r_t + (A^\top - K)(\bar v_{t+1} - \bar V_{t+1}a_t) \tag{42}$$

They correspond exactly to the Recatti equations (26), (27) except for the dependence on Q which interacts directly with the control cost metric H.

Yet another way to write them is:

$$\bar V_t = R_t + A^\top\big[\mathbf{I} - \bar V_{t+1}[\bar V_{t+1} + (Q + BH^{\text{-}1}B^\top)^{\text{-}1}]^{\text{-}1}\big]\bar V_{t+1}A \tag{43}$$
$$\bar v_t = r_t + A^\top\big[\mathbf{I} - \bar V_{t+1}[\bar V_{t+1} + (Q + BH^{\text{-}1}B^\top)^{\text{-}1}]^{\text{-}1}\big](\bar v_{t+1} - \bar V_{t+1}a_t) \tag{44}$$

Proof. Since all factors are pairwise we can use the expression (**??**) for the messages. We have

$\mu_{x_{t\text{-}1}\to x_t}(x_t) = \int_{x_{t\text{-}1}} \text{d}x_{t\text{-}1}\; P(x_t \,|\, x_{t\text{-}1})\; \mu_{x_{t\text{-}2}\to x_{t\text{-}1}}(x_{t\text{-}1})\; \mu_{\hat c_{t\text{-}1}\to x_{t\text{-}1}}(x_{t\text{-}1})$

$= \int_{x_{t\text{-}1}} \text{d}x_{t\text{-}1}\; \mathcal{N}(x_t \,|\, A_{t\text{-}1}x_{t\text{-}1} + a_{t\text{-}1}, Q + B_{t\text{-}1}H^{\text{-}1}B_{t\text{-}1}^\top)\; \mathcal{N}(x_{t\text{-}1} \,|\, s_{t\text{-}1}, S_{t\text{-}1})\; \mathcal{N}[x_{t\text{-}1} \,|\, r_{t\text{-}1}, R_{t\text{-}1}]$

Using the product rule (4) on the last two terms gives a Gaussian $\mathcal{N}(s_{t\text{-}1} \,|\, R_{t\text{-}1}^{\text{-}1}r_{t\text{-}1}, S_{t\text{-}1} + R_{t\text{-}1}^{\text{-}1})$ independent of x_t which we can subsume in the normalization. What remains is

$\mu_{x_{t\text{-}1}\to x_t}(x_t) \propto \int_{x_{t\text{-}1}} \text{d}x_{t\text{-}1}\; \mathcal{N}(x_t \,|\, A_{t\text{-}1}x_{t\text{-}1} + a_{t\text{-}1}, Q + B_{t\text{-}1}H^{\text{-}1}B_{t\text{-}1}^\top)\; \mathcal{N}[x_{t\text{-}1} \,|\, S_{t\text{-}1}^{\text{-}1}s_{t\text{-}1} + r_{t\text{-}1}, S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1}]$

$= \int_{x_{t\text{-}1}} \text{d}x_{t\text{-}1}\; \mathcal{N}(x_t \,|\, A_{t\text{-}1}x_{t\text{-}1} + a_{t\text{-}1}, Q + B_{t\text{-}1}H^{\text{-}1}B_{t\text{-}1}^\top)\; \mathcal{N}(x_{t\text{-}1} \,|\, (S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1}(S_{t\text{-}1}^{\text{-}1}s_{t\text{-}1} + r_{t\text{-}1}), (S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1})$

$= \mathcal{N}(x_t \,|\, A_{t\text{-}1}(S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1}(S_{t\text{-}1}^{\text{-}1} s_{t\text{-}1} + r_{t\text{-}1}) + a_{t\text{-}1}, Q + B_{t\text{-}1} H^{\text{-}1} B_{t\text{-}1}^\top + A_{t\text{-}1}(S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1})^{\text{-}1} A_{t\text{-}1}^\top)$

which gives the messages as in (33). For comparison we also give the canonical representation. Let $\bar S_{t\text{-}1} = S_{t\text{-}1}^{\text{-}1} + R_{t\text{-}1}$ and $\bar s_t = S_{t\text{-}1}^{\text{-}1} s_{t\text{-}1} + r_{t\text{-}1}$,

$$\begin{aligned}
S_t &= Q + B_{t\text{-}1} H^{\text{-}1} B_{t\text{-}1}^\top + A_{t\text{-}1} \bar S_{t\text{-}1}^{\text{-}1} A_{t\text{-}1}^\top \\
&= A_{t\text{-}1}\{A_{t\text{-}1}^{\text{-}1}(Q + B_{t\text{-}1} H^{\text{-}1} B_{t\text{-}1}^\top) A_{t\text{-}1}^{\text{-}\top} + \bar S_{t\text{-}1}^{\text{-}1}\} A_{t\text{-}1}^\top \\
S_t^{\text{-}1} &= A_{t\text{-}1}^{\text{-}\top}\{\bar S_{t\text{-}1} - \bar S_{t\text{-}1}[\bar S_{t\text{-}1} + A_{t\text{-}1}^\top (Q + B_{t\text{-}1} H^{\text{-}1} B_{t\text{-}1}^\top)^{\text{-}1} A_{t\text{-}1}] \bar S_{t\text{-}1}\} A_{t\text{-}1}^{\text{-}1} \\
s_t &= a_{t\text{-}1} + A_{t\text{-}1} \bar S_{t\text{-}1}^{\text{-}1} \bar s_t \\
S_t^{\text{-}1} s_t &= A_{t\text{-}1}^{\text{-}\top} \bar S_{t\text{-}1} A_{t\text{-}1}^{\text{-}1} a_{t\text{-}1} + A_{t\text{-}1}^{\text{-}\top} \bar s_t - A_{t\text{-}1}^{\text{-}\top} \bar S_{t\text{-}1}[\cdots] \bar S_{t\text{-}1} A_{t\text{-}1}^{\text{-}1} a_{t\text{-}1} - A_{t\text{-}1}^{\text{-}\top} \bar S_{t\text{-}1}[\cdots] \bar s_t \\
&= A_{t\text{-}1}^{\text{-}\top}(\bar s_t + \bar S_{t\text{-}1} A_{t\text{-}1}^{\text{-}1} a_{t\text{-}1}) - A_{t\text{-}1}^{\text{-}\top} \bar S_{t\text{-}1}[\cdots](\bar s_t + \bar S_{t\text{-}1} A_{t\text{-}1}^{\text{-}1} a_{t\text{-}1})
\end{aligned}$$

We repeat the derivation for $\mu_{x_{t+1}\to x_t}(x_t)$,

$$\begin{aligned}
\mu_{x_{t+1}\to x_t}(x_t) &= \textstyle\int_{x_{t+1}} dx_{t+1}\, P(x_{t+1} \,|\, x_t)\, \mu_{x_{t+2}\to x_{t+1}}(x_{t+1})\, \mu_{\hat c_{t+1}\to x_{t+1}}(x_{t+1}) \\
&= \textstyle\int_{x_{t+1}} dx_{t+1}\, \mathcal{N}(x_{t+1} \,|\, A_t x_t + a_t, Q + B_t H^{\text{-}1} B_t^\top)\, \mathcal{N}(x_{t+1} \,|\, v_{t+1}, V_{t+1})\, \mathcal{N}[x_{t+1} \,|\, r_{t+1}, R_{t+1}] \\
&\propto \textstyle\int_{x_{t+1}} dx_{t+1}\, \mathcal{N}(x_{t+1} \,|\, A_t x_t + a_t, Q + B_t H^{\text{-}1} B_t^\top)\, \mathcal{N}[x_{t+1} \,|\, V_{t+1}^{\text{-}1} v_{t+1} + r_{t+1}, V_{t+1}^{\text{-}1} + R_{t+1}] \\
&= \mathcal{N}(A_t x_t + a_t \,|\, (V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}(V_{t+1}^{\text{-}1} v_{t+1} + r_{t+1}), Q + B_t H^{\text{-}1} B_t^\top + (V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}) \\
&= \mathcal{N}(x_t \,|\, -A_t^{\text{-}1} a_t + A_t^{\text{-}1}(V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}(V_{t+1}^{\text{-}1} v_{t+1} + r_{t+1}), A_t^{\text{-}1}[Q + B_t H^{\text{-}1} B_t^\top + (V_{t+1}^{\text{-}1} + R_{t+1})^{\text{-}1}] A_t^{\text{-}\top})
\end{aligned}$$

For this backward message it is instructive to derive the canonical representation. Let $\bar V_{t+1} = V_{t+1}^{\text{-}1} + R_{t+1}$ and $\bar v_{t+1} = V_{t+1}^{\text{-}1} v_{t+1} + r_{t+1}$,

$$\begin{aligned}
V_t^{\text{-}1} &= A^\top [\bar V_{t+1}^{\text{-}1} + Q + B H^{\text{-}1} B^\top]^{\text{-}1} A \\
&= A^\top \bar V_{t+1} A - A^\top \bar V_{t+1}[\bar V_{t+1} + (Q + B H^{\text{-}1} B^\top)^{\text{-}1}]^{\text{-}1} \bar V_{t+1} A \\
&= A^\top \{\bar V_{t+1} - \bar V_{t+1}[\bar V_{t+1} + (Q + B H^{\text{-}1} B^\top)^{\text{-}1}]^{\text{-}1} \bar V_{t+1}\} A \\
v_t &= -A_t^{\text{-}1} a_t + A_t^{\text{-}1} \bar V_{t+1}^{\text{-}1} \bar v_{t+1} \\
V_t^{\text{-}1} v_t &= -A^\top \bar V_{t+1} a_t + A^\top \bar v_{t+1} + A^\top \bar V_{t+1}[\cdots]^{\text{-}1} \bar V_{t+1} a_t - A^\top \bar V_{t+1}[\cdots]^{\text{-}1} \bar v_{t+1} \\
&= A^\top(\bar v_{t+1} - \bar V_{t+1} a_t) - A^\top \bar V_{t+1}[\cdots]^{\text{-}1}(\bar v_{t+1} - \bar V_{t+1} a_t)
\end{aligned}$$

■

2.3 Special case: kinematic control

We assume a kinematic control problem: (We write q instead of x.) The Process is simply

$$q_{t+1} = q_t + u_t + \xi \,. \tag{45}$$

This means we have

$$A_t = B_t = 1 \,, \quad a_t = 0 \tag{46}$$

2.4 Special case: multiple kinematic task variables

Let $\phi_i : q \mapsto x_i$ be a kinematic mapping to a task variable y_i and $J_i(q)$ its Jacobian. We assume we are given targets $y^*_{i,0:T}$ in the task space and (time-dependent) precisions $\varrho_{i,t}$ by which we want to follow the task targets. We have

$$\begin{aligned}
c_t(q_t, u_t) &= \|u_t\|_H + \sum_{i=1}^m \|y^*_{i,t} - \phi_i(q_t)\|_{C_i^{\text{-}1}} + \\
&\approx \|u_t\|_H + \sum_{i=1}^m \|y^*_{i,t} - \phi_i(\hat q_t) + J_i \hat q_t - J_i q_t\|_{C_i^{\text{-}1}} \,, \quad J_i = J_i(\hat q_t) \\
&= \|u_t\|_H + \sum_{i=1}^m q_t^\top J_i^\top C_i^{\text{-}1} J_i q_t - 2(y^*_{i,t} - \phi_i(\hat q_t) + J_i \hat q_t)^\top C_i^{\text{-}1} J_i q_t + \text{const} \qquad (47) \\
R_t &= \sum_{i=1}^m J_i^\top C_i^{\text{-}1} J_i \qquad (48) \\
r_t &= \sum_{i=1}^m J_i^\top C_i^{\text{-}1}(y^*_{i,t} - \phi_i(\hat q_t) + J_i \hat q_t) \qquad (49)
\end{aligned}$$

Note: the product of a fwd message with a task message corresponds to the classical optimal control for multiple regularized task variables (92)

$$\mathcal{N}(q_t \,|\, s_t, S_t)\, \mathcal{N}[q_t \,|\, r_t, R_t] \propto \mathcal{N}(q_t \,|\, b, B) \tag{50}$$

$$b = \Big[R_t + S_t^{\text{-}1}\Big]^{\text{-}1} \Big[r_t + S_t^{\text{-}1} s_t\Big] \tag{51}$$

2.5 Special case: Pseudo-dynamic process

We replace x by $\bar q_t = \begin{pmatrix} q_t \\ \dot q_t \end{pmatrix}$ and assume u_t corresponds directly to accelerations:

$$P(q_{t+1} \mid \dot q_t, q_t) = \mathcal{N}(q_{t+1} \mid q_t + \tau \dot q_{t+1}, W^{-1}) \;, \tag{52}$$

$$P(\dot q_{t+1} \mid \dot q_t, u_t) = \mathcal{N}(\dot q_{t+1} \mid \dot q_t + \tau u_t, Q) \;, \tag{53}$$

$$\begin{pmatrix} q_{t+1} \\ \dot q_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q_t \\ \dot q_t \end{pmatrix} + \begin{pmatrix} \tau^2 \\ \tau \end{pmatrix} u_t + \xi \;, \quad \langle d\xi d\xi^\top \rangle = \begin{pmatrix} W^{-1} & 0 \\ 0 & Q \end{pmatrix} \tag{54}$$

$$A = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} , \quad B = \begin{pmatrix} \tau^2 \\ 1 \end{pmatrix} , \quad a = 0 \tag{55}$$

The following equations might be computationally more efficient than the general Ricatti recursion – but I'm not sure,

$$u^*_t(x) = -(H + B^\top V_{t+1} B)^{-1} B^\top (V_{t+1}(Ax + a) - v_{t+1})$$

$$V_t = \begin{pmatrix} V^1_t & V^2_t \\ V^3_t & V^4_t \end{pmatrix} = R + V_\cdot + \tau \begin{pmatrix} 0 & V^1_\cdot \\ V^1_\cdot & V^2_\cdot + V^3_\cdot + \tau V^1_\cdot \end{pmatrix} - \tau^2 \begin{pmatrix} V^3_\cdot \\ \tau V^3_\cdot + V^4_\cdot \end{pmatrix} (H + \tau^2 V^4_\cdot)^{-1} (V^3_\cdot \;,\; \tau V^3_\cdot + V^4_\cdot) \;, \tag{56}$$

$$v_t = \begin{pmatrix} v^1_t \\ v^2_t \end{pmatrix} = r + v_\cdot + \tau \begin{pmatrix} 0 \\ v^1 \end{pmatrix} - \tau^2 \begin{pmatrix} V^3_\cdot \\ \tau V^3_\cdot + V^4_\cdot \end{pmatrix} (H + \tau^2 V^4_\cdot)^{-1} v_\cdot \tag{57}$$

$$u^*_t(x) = -(H + \tau^2 V^4_\cdot)^{-1} (V_{t+1} A x - v_{t+1}) \tag{58}$$

2.6 Special case: dynamic process

We replace x by $\bar q_t = \begin{pmatrix} q_t \\ \dot q_t \end{pmatrix}$

$$P(q_{t+1} \mid \dot q_t, q_t) = \mathcal{N}(q_{t+1} \mid q_t + \tau \dot q_{t+1}, W^{-1}) \;, \tag{59}$$

$$P(\dot q_{t+1} \mid \dot q_t, u_t) = \mathcal{N}(\dot q_{t+1} \mid \dot q_t + \tau M^{-1}(u_t + F), Q) \;, \tag{60}$$

$$\begin{pmatrix} q_{t+1} \\ \dot q_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q_t \\ \dot q_t \end{pmatrix} + \begin{pmatrix} \tau^2 \\ \tau \end{pmatrix} M^{-1}(u_t + F) + \xi \;, \quad \langle d\xi d\xi^\top \rangle = \begin{pmatrix} W^{-1} & 0 \\ 0 & Q \end{pmatrix} \tag{61}$$

$$A = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} , \quad B = \begin{pmatrix} \tau^2 M^{-1} \\ \tau M^{-1} \end{pmatrix} , \quad a = \begin{pmatrix} \tau^2 M^{-1} F \\ \tau M^{-1} F \end{pmatrix} \tag{62}$$

The following equations might be computationally more efficient than the general Ricatti recursion – but I'm not sure,

$$u^*_t(x) = -(H + B^\top V_{t+1} B)^{-1} B^\top (V_{t+1}(Ax + a) - v_{t+1})$$

$$V_t = \begin{pmatrix} V^1_t & V^2_t \\ V^3_t & V^4_t \end{pmatrix} = R + V_\cdot + \tau \begin{pmatrix} 0 & V^1_\cdot \\ V^1_\cdot & V^2_\cdot + V^3_\cdot + \tau V^1_\cdot \end{pmatrix} - \tau^2 \begin{pmatrix} V^3_\cdot \\ \tau V^3_\cdot + V^4_\cdot \end{pmatrix} M^{-1} (H + \tau^2 M^{-\top} V^4_\cdot M^{-1})^{-1} M^{-\top} (V^3_\cdot \;,\; \tau V^3_\cdot + V^4_\cdot) \;, \tag{63}$$

$$v_t = \begin{pmatrix} v^1_t \\ v^2_t \end{pmatrix} = r + v_\cdot + \tau \begin{pmatrix} 2 V^2_\cdot F \\ v^1 + 2 V^4_\cdot F + 2\tau V^2_\cdot F \end{pmatrix} - \tau^2 \begin{pmatrix} V^3_\cdot \\ \tau V^3_\cdot + V^4_\cdot \end{pmatrix} M^{-1} (H + \tau^2 M^{-\top} V^4_\cdot M^{-1})^{-1} M^{-\top} (v_\cdot + 2\tau \begin{pmatrix} V^2_\cdot F \\ V^4_\cdot F \end{pmatrix}) \tag{64}$$

$$u^*_t(x) = -(H + \tau^2 M^{-\top} V^4_\cdot M^{-1})^{-1} M^{-\top} (V_{t+1}(Ax + a) - v_{t+1}) \tag{65}$$

2.7 Special case: multiple dynamic/kinematic task variables

As before we have access to kinematic functions $\phi_i(q)$ and Jacobians $J_i(q)$. We are given task targets $x^*_{i,0:T}$ and $\dot x^*_{i,0:T}$ and want to follow them with (time-dependent) precisions $\varrho_{i,0:T}$ and $\nu_{i,0:T}$. We have

$$\begin{aligned} c(q_t, \dot q_t, u_t) &= \sum_{i=1}^m \varrho_{i,t} [x^*_{i,t} - \phi_i(q_t)]^2 + \nu_{i,t} [\dot x^*_{i,t} - J_i \dot q_t]^2 + u_t^\top H_t u_t \\ &\approx \sum_{i=1}^m \varrho_{i,t} [q_t^\top J_i^\top J_i q_t - 2(x^*_{i,t} - \phi_i(\hat q_t) + J_i \hat q_t)^\top J_i q_t + \mathit{const}] + \nu_{i,t} [\dot q_t^\top J_i^\top J_i \dot q_t - 2(\dot x^*_{i,t})^\top J_i \dot q_t + \mathit{const}] + u_t^\top H_t u_t \end{aligned} \tag{66}$$

$$R_t = \sum_{i=1}^m \begin{pmatrix} \varrho_{i,t} J_i^\top J_i & 0 \\ 0 & \nu_{i,t} J_i^\top J_i \end{pmatrix} \tag{67}$$

$$r_t = \sum_{i=1}^m \begin{pmatrix} \varrho_{i,t} J_i^\top (x^*_{i,t} - \phi_i(\hat q_t) + J_i \hat q_t) \\ \nu_{i,t} J_i^\top \dot x^*_{i,t} \end{pmatrix} \tag{68}$$

3 The optimization view on classical control laws

In this section we will first review classical control laws as minimization of a basic loss function. Since this loss function has a Bayesian interpretation it will be straight-forward to develop also a Bayesian view on these control laws. The Bayesian inference approach can then be generalized to what we actually aim for: motion planning in temporal probabilistic models.

3.1 General quadratic loss function and constraints

Let $y \in \mathbb{R}^d$ and $q \in \mathbb{R}^n$. Given y, consider the problem of finding q that minimizes

$$L = \|q\|_W + \|y - Jq\|_{C^{-1}} - 2h^\top W q \tag{69}$$

where $\|q\|_W = q^\top W q$ denotes a norm and C and W are symmetric positive definite matrices. This loss function can be interpreted as follows: The first term "measures" how well a constraint $y = Jq$ is fulfilled relative to a covariance matrix C, the second term "measures" $q = 0$ with metric W, the third term "measures" the scalar product between q and h w.r.t. W.

The solution can easily be found by taking the derivative

$$\frac{\partial L}{\partial q} = 2q^\top W - 2(y - Jq)^\top C^{-1} J - 2h^\top W$$
$$\frac{\partial L}{\partial q} = 0 \quad \Rightarrow \quad q^* = (J^\top C^{-1} J + W)^{-1} (J^\top C^{-1} y + W h)$$

Using the Woddbury and related identities and definitions as given in section 1 we can rewrite the solution in several forms:

$$\begin{aligned} q^* &= (J^\top C^{-1} J + W)^{-1} (J^\top C^{-1} y + W h) && (70)\\ &= (J^\top C^{-1} J + W)^{-1}\ J^\top C^{-1}\ y\ +\ [\mathbf{I}_n - (J^\top C^{-1} J + W)^{-1}\ J^\top C^{-1} J]\ h && (71)\\ &= (J^\top C^{-1} J + W)^{-1}\ J^\top C^{-1}\ (y - Jh)\ +\ h && (72)\\ &= W^{-1} J^\top\ (J W^{-1} J^\top + C)^{-1}\ y\ +\ [\mathbf{I}_n - W^{-1} J^\top\ (J W^{-1} J^\top + C)^{-1} J]\ h && (73)\\ &= W^{-1} J^\top\ (J W^{-1} J^\top + C)^{-1}\ (y - Jh)\ +\ h\ . && (74) \end{aligned}$$

This also allows us to properly derive the following limits:

$$\begin{aligned} C \to 0 :&\quad q^* = J^\sharp_W y + (\mathbf{I}_n - J^\sharp_W J)\ h\ = J^\sharp_W (y - Jh) + h && (75)\\ W \to 0 :&\quad q^* = J^\natural_C y + (\mathbf{I}_n - J^\natural_C J)\ h\ = J^\natural_C (y - Jh) + h \\ W = \lambda \mathbf{I}_n :&\quad q^* = J^\top (J J^\top + \lambda C)^{-1} y\ + [\mathbf{I}_n - J^\top (J J^\top + \lambda C)^{-1} J]\ h && (76)\\ C = \sigma \mathbf{I}_d :&\quad q^* = (J^\top J + \sigma W)^{-1} J^\top y + [\mathbf{I}_n - (J^\top J + \sigma W)^{-1} J^\top J]\ h \end{aligned}$$

These limits can be interpreted as follows. $C \to 0$: we need to fulfill the constraint $y = Jq$ exactly. $C = \sigma \mathbf{I}_d$: we use a standard squared error measure for $y \approx Jq$. $W \to 0$: we do not care about the norm $\|q\|_W$ (i.e., no regularization); but interestingly, the cost term $h^\top W q$ has a nullspace effect also in this limit. $W = \lambda \mathbf{I}_n$: we use a standard ridge as regulariser.

The first of these limits is perhaps the most important. It corresponds to a hard constraint, that is, (75) is the solution to

$$\operatorname*{argmin}_q \Big[\|q\|_W - 2h^\top W q\Big] \qquad \text{such that} \qquad y = Jq \tag{77}$$

The loss function (69) has many applications, as we discuss in the following.

3.2 Ridge regression

Let us first give an "off topic" example from machine learning: In ridge regression, when we have d samples of n-dimensional inputs and 1D outputs, we have a minimization problem

$$L = \|y - X\beta\| + \lambda\|\beta\|$$

with a input data matrix $X \in \mathbb{R}^{d\times n}$, an output data vector $y \in \mathbb{R}^d$ and a regressor $\beta \in \mathbb{R}^n$. The first term measures the standard squared error (with uniform output covariance $C = \mathbf{I}_d$), the second is a regulariser (or stabilizer) as introduced by Tikhonov. The special form $\lambda\|\beta\|$ of the regulariser is called ridge. The solution is given by equation (76) when replacing the notation according to $q \rightsquigarrow \beta$, $y \rightsquigarrow y$, $J \rightsquigarrow X$, $C \rightsquigarrow \mathbf{I}_d$, $W \rightsquigarrow \lambda\mathbf{I}_n$, $h \rightsquigarrow 0$:

$$\beta^* = X^\top(XX^\top + \lambda\mathbf{I}_d)^{\text{-}1}y \;.$$

In the Bayesian interpretation of ridge regression, the ridge $\lambda\|\beta\|$ defines a prior $\propto \exp\{-\frac{1}{2}\lambda\|\beta\|\}$ over the regressor β. The above equation gives the MAP β. Since ridge regression has a Bayesian interpretation, also standard motion rate control, as discussed shortly, will have a Bayesian interpretation.

3.3 Motion rate control (pseudo-inverse kinematics)

Consider a robot with n DoFs and a d-dimensional task space with $d < n$ (e.g., an endeffector state). The current joint state is $q \in \mathbb{R}^n$. In a given state we can compute the end-effector Jacobian J and we are given a joint space potential $H(q)$. We would like to compute joint velocities $\dot q$ which fulfill the task constraint $\dot y = J\dot q$ while minimizing the absolute joint velocity $\|\dot q\|_W$ and following the negative gradient $h = -W^{\text{-}1}\nabla H(q)$. In summary, the problem and its solution are

$$\text{(problem)}\quad \dot q^* = \underset{\dot q}{\text{argmin}}\left[\|\dot q\|_W + 2\nabla H(q)^\top\dot q\right] \qquad \text{such that} \qquad J\dot q = \dot y \tag{78}$$

$$\text{(solution)}\quad \dot q^* = J^\sharp_W\dot y - (\mathbf{I}_n - J^\sharp_W J)\, W^{\text{-}1}\nabla H(q) \;. \tag{79}$$

The solution was taken from (75) by replacing the notation according to $q \rightsquigarrow \dot q$, $y \rightsquigarrow \dot y$.

Note that we have derived pseudo-inverse kinematics from a basic constrained quadratic optimization problem.

Let us repeat this briefly for case when time is discretized. We can formulate the problem as

$$q_t^* = \underset{q_t}{\text{argmin}}\left[\|q_t - q_{t\text{-}1}\|_W - 2h^\top Wq_t\right] \qquad \text{such that} \qquad \phi(q_t) = y_t \tag{80}$$

Generally, the constraint $\phi(q_t) = y_t$ is non-linear. We linearize it at $q_{t\text{-}1}$ and get the simpler problem and its solution

$$\text{(problem)}\quad q_t^* = \underset{q_t}{\text{argmin}}\left[\|q_t - q_{t\text{-}1}\|_W - 2h^\top Wq_t\right] \qquad \text{such that} \qquad J(q_t - q_{t\text{-}1}) = y_t - \phi(q_{t\text{-}1}) \tag{81}$$

$$\text{(solution)}\quad q_t^* = q_{t\text{-}1} + J^\sharp_W[y_t - \phi(q_{t\text{-}1})] + (\mathbf{I}_n - J^\sharp_W J)\; h \;. \tag{82}$$

The solution was taken from (75) by replacing the notation according to $q \rightsquigarrow (q_t - q_{t\text{-}1})$, $y \rightsquigarrow (y_t - \phi(q_{t\text{-}1}))$.

3.4 Regularized inverse kinematics (singularity robust motion rate control)

Under some conditions motion rate control is infeable, for instance when the arm cannot be further streched to reach a desired endeffector position. In this case the computation of the pseudo-inverse $J^\sharp_W$ becomes singular. Classical control developed the singularity robuse pseudo-inverse (Nakamura & Hanafusa, 1986), which can be interpreted as regularizing the computation of the pseudo-inverse, or as relaxing the hard task constraint. In our framework this corresponds to *not* taking the limit $C \to 0$. This regularized inverse kinematics is given as

$$\text{(problem)}\quad \dot q^* = \underset{\dot q}{\text{argmin}}\left[\|\dot q\|_W + \|\dot y - J\dot q\|_{C^{\text{-}1}} - 2h^\top W\dot q\right] \tag{83}$$

$$\text{(solution)}\quad \dot q^* = J^\sharp_{WC}\dot y + (\mathbf{I}_n - J^\sharp_{WC}J)\; h \;. \tag{84}$$

$$J^\sharp_{WC} := (J^\top C^{\text{-}1}J + W)^{\text{-}1}\; J^\top C^{\text{-}1} = W^{\text{-}1}J^\top\,(JW^{\text{-}1}J^\top + C)^{\text{-}1} \tag{85}$$

The solution was taken from (71) by replacing the notation according to $q \rightsquigarrow \dot q$, $y \rightsquigarrow \dot y$. Note that $J^\sharp_{WC}$ is a regularization of $J^\sharp_W$ (defined in (11)). Equations (70-74) give many interesting alternatives to write this control law.

The linearized ($\hat\phi$ is linearized at $q_{t\text{-}1}$ as above) time discretized version is:

$$\text{(problem)}\quad q_t^* = \underset{q_t}{\text{argmin}}\left[\|q_t - q_{t\text{-}1}\|_W + \|y_t - \hat\phi(q_t)\|_{C^{\text{-}1}} - 2h^\top Wq_t\right] \tag{86}$$

$$\text{(solution)}\quad q_t^* = q_{t\text{-}1} + J^\sharp_{WC}[y_t - \phi(q_{t\text{-}1})] + (\mathbf{I}_n - J^\sharp_{WC}J)\; h \;. \tag{87}$$

3.5 Multiple regularized task variables

Assume we have m task variables $y_1, .., y_m$, where the ith variable $y_i \in \mathbb{R}^{d_i}$ is d_i-dimensional. Also assume that we regularize w.r.t. each task variable, that is, we have different error metrics C_i^{-1} in each task space. We want to follow all of the tasks and express this as the optimization problem and its solution

$$\text{(problem)} \quad \dot q^* = \underset{\dot q}{\text{argmin}} \Big[\|\dot q\|_W + \sum_{i=1}^m \|\dot y_i - J\dot q\|_{C_i^{-1}} - 2h^\top W \dot q \Big] \tag{88}$$

$$\text{(solution)} \quad \dot q^* = \Big[\sum_{i=1}^m J^\top C_i^{-1} J + W\Big]^{-1} \Big[\sum_{i=1}^m J^\top C_i^{-1} \dot y_i + Wh\Big] . \tag{89}$$

The solution was taken from (70) in the following way: We can collect all task variables into one bit task vector

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} , \quad J = \begin{pmatrix} J_1 \\ \vdots \\ J_m \end{pmatrix} , \quad C = \begin{pmatrix} C_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & C_m \end{pmatrix} \quad \Rightarrow \quad J^\top C^{-1} J = \sum_{i=1}^m J_i^\top C_i^{-1} J_i \,, \quad J^\top C^{-1} \dot y = \sum_{i=1}^m J_i^\top C_i^{-1} \dot y_i \tag{90}$$

And the linearized time discretized version:

$$\text{(problem)} \quad q_t^* = \underset{q_t}{\text{argmin}} \Big[\|q_t - q_{t\text{-}1}\|_W + \sum_{i=1}^m \|y_{i,t} - \hat\phi_i(q_t)\|_{C_i^{-1}} - 2h^\top W \dot q \Big] \tag{91}$$

$$\text{(solution)} \quad q_t^* = q_{t\text{-}1} + \Big[\sum_{i=1}^m J^\top C_i^{-1} J + W\Big]^{-1} \Big[\sum_{i=1}^m J^\top C_i^{-1} [y_{i,t} - \phi(q_{t\text{-}1})] + Wh\Big] . \tag{92}$$

3.6 Multiple prioritized task variables (prioritized inverse kinematics)

The case of multiple task variables is classically not addressed by regularizing all of them, but by imposing a hierarchy on them (Nakamura et al., 1987; Baerlocher & Boulic, 2004). Let us first explain the classical prioritized inverse kinematics: The control law is based on standard motion rate control but iteratively projects each desired task rate $\dot y_i$ in the remaining nullspace of all higher level control signals. Initializing the nullspace projection with $N_0 = \mathbf{I}$ and $\dot q_0 = 0$, the control law is defined by iterating for $i = 1, .., m$

$$\bar J_i = J_i N_{i\text{-}1} \,, \quad \dot q_i = \dot q_{i\text{-}1} + \bar J_i^\sharp (\dot y_i - J_i \, \dot q_{i\text{-}1}) \,, \quad N_i = N_{i\text{-}1} - \bar J_i^\sharp \bar J_i \,. \tag{93}$$

We call $\bar J_i$ a nullspace Jacobian which has the property that $\bar J_i^\sharp$ projects to changes in q that do not change control variables $x_{1,..,i\text{-}1}$ with higher priority. Also an additional nullspace movement h in the remaining nullspace of all control signals can be included when defining the final control law as

$$\dot q^* = \dot q_m + N_m h \,. \tag{94}$$

In effect, the first task rate $\dot x_1$ is guaranteed to be fulfilled exactly. The second $\dot x_2$ is guaranteed to be fulfilled "as best as possible" given that $\dot x_1$ must be fulfilled, et cetera.

This hierarchical projection of task can also be derived by starting with the regularize task variables as in the problem (88) and then iteratively taking the limit $C_i \to 0$ starting with $i = 1$ up to $i = m$. More formally, the iterative limit corresponds to $C_i = \epsilon^{m-i} \mathbf{I}_{d_i}$ and $\epsilon \to 0$. For $m = 2$ task variables one can prove the equivalence between prioritized inverse kinematics and the hierarchical classical limit of the MAP motion exactly (by directly applying the Woodbury identity). For $m > 2$ we could not find an elegant proof but we numerically confirmed this limit for up to $m = 4$.

Non-zero task variances can again be interpreted as regularizers. Note that without regularizers the standard prioritized inverse kinematics is numerically brittle. Handling many control signals (e.g., the over-determined case $\sum d_i > n$) is problematic since the nullspace-projected Jacobians will become singular (with rank $< d_i$). For non-zero regularizations C_i the computations in equation (92) are numerically robust.

3.7 Optimal dynamic control (incl. operational space control)

Consider a robot with dynamics $\ddot q = M^{-1}(u + F)$, where M is some generalized mass matrix, F subsumes external (also Coriolis and gravitational) forces, and u is the n-dimensional torque control signal. We want to compute a control signal u which generates an acceleration such that a general task constraint $\ddot y = J\ddot q + \dot J \dot q$ remains fulfilled while also minimizing the absolute norm $\|u\|_H$ of the control. The problem and its solution can be written as

$$\text{(problem)} \quad u^* = \underset{u}{\operatorname{argmin}} \Big[\|u\|_H - 2h^\top H u \Big] \qquad \text{such that} \qquad \ddot y - \dot J \dot q - JM^{\text{-}1}(u + F) = 0 \tag{95}$$

$$\text{(solution)} \quad u^* = T^\sharp_H(\ddot y - \dot J \dot q - TF) + (\mathbf{I}_n - T^\sharp_H T)\ h\ , \quad \text{with } T = JM^{\text{-}1}\ . \tag{96}$$

The solution was taken from (75) by replacing the notation according to $q \rightsquigarrow u$, $y \rightsquigarrow (\ddot y - \dot J \dot q - JM^{\text{-}1}F)$, $J \rightsquigarrow JM^{\text{-}1}$. For $h = 0$ this solution is identical to Theorem 1 in (Peters et al., 2005). Peters et al. discuss in detail stability issues and important special cases of this control scheme. A common special case is $H = M^{\text{-}1}$, which is called operational space control.

References

Baerlocher, P., & Boulic, R. (2004). An inverse kinematic architecture enforcing an arbitrary number of strict priority levels. *The Visual Computer*.

Nakamura, Y., & Hanafusa, H. (1986). Inverse kinematic solutions with singularity robustness for robot manipulator control. *Journal of Dynamic Systems, Measurement and Control*, *108*.

Nakamura, Y., Hanafusa, H., & Yoshikawa, T. (1987). Task-priority based redundancy control of robot manipulators. *Int. Journal of Robotics Research*, *6*.

Peters, J., Mistry, M., Udwadia, F. E., Cory, R., Nakanishi, J., & Schaal, S. (2005). A unifying framework for the control of robotics systems. *IEEE Int. Conf. on Intelligent Robots and Systems (IROS 2005)* (pp. 1824–1831).