Duality between cost and density in optimal control

Rantzer, Anders; Hedlund, Sven

Published in:
Proceedings. 42nd IEEE Conference on Decision and Control, 2003

DOI:
10.1109/CDC.2003.1272774

2003

Proceedings of the 42nd IEEE
Conference on Decision and Control
Maui, Hawaii USA, December 2003

Duality between cost and density in optimal control

Anders Rantzer and Sven Hedlund

Department of Automatic Control, LTH, Lund University
Box 118, S-221 00 Lund, Sweden
Email: rantzer at control.lth.se

Abstract A theorem on duality between cost functions and density functions in optimal control is derived using the Hahn-Banach theorem. The result puts focus on convexity aspects in control synthesis and the recent theory of almost global stability. In particular, it gives a new proof that existence of a density function is both necessary and sufficient for almost global stability in a nonlinear system.

Keywords Density function, duality, stability, optimal control

1. Introduction

The idea of duality between cost and flow has old roots. In fact, a non-linear problem of optimal transportation stated by G. Monge in 1781 was converted into convex optimization by [Kantorovich, 1942] and inspired much of the later developments in the theory of convex duality. See [Rachev and Rüschendorf, 1998]. Kantorovich later received the Nobel price for related work in mathematical economics.

The ideas were introduced in the context of optimal control by [Young, 1969] using the concept of generalized flow. For later work, see [Vinter, 1993]. More recently, [Rantzer, 2001] introduced the concept of density function as a tool for verification of almost global stability in non-linear systems. The relation to duality theory was then briefly discussed.

The new stability concept has a remarkable convexity property in the context of control synthesis. This was explored for numerical computations in [Rantzer and Parrilo, 2000] and for smooth transitions between different nonlinear controllers in [Rantzer and Ceragioli, 2001].

The purpose of the present paper is to establish the duality between cost functions and density functions in a more rigorous manner. The main result is stated and discussed in section 2. The next section is devoted to the proof. The construction of control law from density functions is described and the main duality argument is given.

2. Main result

Let $f_i \in C^1(\mathbf{R}^n, \mathbf{R}^n)$, $l_i \in C(\mathbf{R}^n, \mathbf{R}^m)$ with $l_i \geq 0$ for $i = 1, \ldots, M$. Let $\Gamma, X \subset \mathbf{R}^n$ be open bounded sets with C^1 boundary and $\overline{\Gamma} \subset X$. Introduce $\mathcal{U}$ as the set of all $(u_1, \ldots, u_M) \in C^1(X, \mathbf{R}^M)$ with $u_i(x) \geq 0$ and $u_1(x) + \cdots + u_M(x) \equiv 1$ such that every solution of $\dot{x} = \sum_i u_i(x) f_i(x)$ starting in X at $t = 0$ stays in X for all $t > 0$. Let the solution with $x(0) = x_0$ be denoted $\phi_u(x_0, t)$. Define

$$V_u(x) = \sum_i \int_0^\infty u_i\left(\phi_u(x,t)\right) l_i\left(\phi_u(x,t)\right) dt$$
$$V^*(x) = \inf_{u \in \mathcal{U}} V_u(x)$$

The main theorem can now be stated as follows:

THEOREM 1
Consider X, f_i, l_i, $\mathcal{U}$, ϕ_u and V^* as above. Let $l_i > 0$ outside $\overline{\Gamma}$ and $l_i = 0$ inside. Define $\psi \in C(\overline{X})$ with $\psi > 0$. Then

$$\int_X \psi(x) V^*(x) dx = \inf_{\rho_i} \sum_{i=1}^M \int_X l_i(x) \rho_i(x) dx \qquad (1)$$

where inf is taken over $\rho_i \in C_0^1(\overline{X})$ with $\rho_i > 0$ in X and

$$\sum_{i=1}^M \nabla \cdot (f_i(x) \rho_i(x)) > \psi(x) \qquad x \in X \setminus \Gamma \qquad (2)$$

Moreover, for all ρ_i satisfying the conditions above, $u := (\rho_1, \ldots, \rho_M)/(\sum_i \rho_i)$ is an element in $\mathcal{U}$ and

$$\int_X \psi(x) V_u(x) dx < \sum_{i=1}^M \int_X l_i(x) \rho_i(x) dx$$

□

Before giving the proof in a later section, we make a few

remarks.

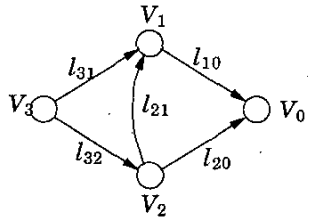

Figure 1: The products produced in nodes 1-3 should be transported to the consumer in node 0 while minimizing the transportation cost.

The case of no control variable (M=1) In this case, the value of the integral

$$\int_X \psi(x)V^*(x)dx \tag{3}$$

is interesting as a stability indicator. A finite value of the integral means that $V^*(x) = \int_0^\infty l(x(t))dt$ is finite for almost all $x \in X \setminus \Gamma$. For all these initial states, the trajectory must approach Γ as $t \to \infty$. Hence, Theorem 1 proves existence of non-negative $\rho \in C_0^1(\overline{X})$ such that

$$\nabla \cdot (f(x)\rho(x)) > \psi(x) > 0 \qquad x \in X \setminus \Gamma$$

Conversely, if such a ρ exists, then the theorem shows that (3) is finite and almost all trajectories eventually approach Γ.

Control synthesis by convex optimization It should be noted that the minimization corresponding to the infimum expression in Theorem 1 is a problem of convex optimization. In fact, every multiple $(\rho_1, \ldots, \rho_M)$ that solves the divergence inequality not only gives an upper bound on the optimal value $\int_X \psi(x)V^*(x)dx$, but also corresponds to a control law achieving this upper bound. This can be viewed as the reason behind the previously mentioned convexity property of density functions in control synthesis.

Comparison to a discrete transportation problem It is natural to compare Theorem 1 to the standard linear programming solution to the discrete transportation problem illustrated in Figure 1. Such problems have been studied extensively since the 1940's [Hitchcock, 1941; Ford and Fulkerson, 1962]. Some product is produced with unit rate in each of the three nodes 1-3 and is consumed in node 0. The cost for shipping the product between node i and node j is given by the number l_{ij}. It is well known that the minimal total transportation cost can be found by solving a linear programming problem:

$$\begin{array}{ll} \text{Maximize} & V_1 + V_2 + V_3 - 3V_0 \\ \text{subject to} & V_3 - V_1 \le l_{31} \\ & V_3 - V_2 \le l_{32} \\ & \vdots \\ & V_2 - V_0 \le l_{20} \end{array}$$

Note that there is one variable V_i for each node and one inequality constraint for each path connecting two nodes. For every solution to the inequality constraints, the number $V_i - V_0$ provides a lower bound on the cost for shipping products with unit rate from node i to node 0. The expression $V_1 + V_2 + V_3 - 3V_0$ therefore gives a lower bound on the total transportation cost.

A dual LP problem can be stated as follows.

$$\begin{array}{ll} \text{Minimize} & l_{31}\rho_{31} + l_{32}\rho_{32} + l_{21}\rho_{21} + l_{10}\rho_{10} + l_{20}\rho_{20} \\ \text{subject to} & \rho_{31}, \ldots, \rho_{20} \ge 0 \\ & \rho_{31} + \rho_{32} \ge 1 \\ & -\rho_{31} - \rho_{21} + \rho_{10} \ge 1 \\ & -\rho_{32} + \rho_{21} + \rho_{20} \ge 1 \end{array}$$

For each path connecting two nodes, the variable ρ_{ij} can be interpreted as the transportation density from node i to node j. There is one constraint for each node stating that the total production in this node is at least as big as the assigned value.

3. From densities to control law

The following lemma is essential for Theorem 1:

LEMMA 1
Let $X \subset \mathbf{R}^n$ be open. Given $f \in C^1(X, \mathbf{R}^n)$, suppose that X is invariant to the dynamics $\dot{x} = f(x)$. Let $l \in C(X)$ be strictly positive outside $\overline{\Gamma}$ and zero inside. Let $\rho \in C_0^1(X)$ be non-negative and $(\nabla \cdot (f\rho)(x) > 0$ for $x \in \overline{X} \setminus \Gamma$. Define $V^*(x) = \int_0^\infty l\left(\phi(x,t)\right) dt$. Then

$$\int_X V^*(x)(\nabla \cdot (f\rho)(x)dx = \int_X l(x)\rho(x)dx$$

□

The proof of Lemma 1 uses a version of Liouville's theorem [Rantzer, 2001]:

PROPOSITION 1
Let $f \in \mathbf{C}^1(X, \mathbf{R}^n)$ and let $\rho \in \mathbf{C}^1(X)$ be integrable. For a measurable set Z, assume that $\phi_\tau(Z) = \{\phi_\tau(x) \mid x \in Z\}$ is a subset of X for all τ between 0 and t. Then

$$\int_{\phi(Z,t)} \rho(x)dx - \int_Z \rho(z)dz = \int_0^t \int_{\phi(Z,\tau)} [\nabla \cdot (f\rho)](x)dxd\tau$$

□

Proof of Lemma 1. Consider $T > 0$ and a piecewise constant $l(x) = \sum_i l_i \chi_i(x)$, where χ_i is the characteristic function of the set $X_i \subset X$. Then

$$\begin{aligned}
&\int_X l(x)\rho(x)dx - \sum_i l_i \int_{\phi(X_i,-T)} \rho(x)dx \\
&= \sum_i l_i \left(\int_{X_i} \rho(x)dx - \int_{\phi(X_i,-T)} \rho(x)dx \right) \\
&= \sum_i l_i \int_0^T \int_{\phi(X_i,-t)} \nabla \cdot (f\rho)(z)dzdt \\
&= \sum_i l_i \int_0^T \int_{\phi(X_i,-t)} \nabla \cdot (f\rho)(z)dzdt \\
&= \sum_i l_i \int_0^T \int_X \chi_i(\phi(z,t)) \nabla \cdot (f\rho)(z)dzdt \\
&= \int_X \left(\int_0^T l(\phi(z,t))dt \right) \nabla \cdot (f\rho)(z)dz
\end{aligned}$$

In the limit as $T \to \infty$, this gives

$$\begin{aligned}
&\int_X l(x)\rho(x)dx - \lim_{T\to\infty} \sum_i l_i \int_{\phi(X_i,-T)} \rho(x)dx \\
&= \int_X V^*(x)(\nabla \cdot (f\rho)(x)dx
\end{aligned}$$

If the right hand expression is infinite, then also $\int l(x)\rho(x)dx$ must be infinite, so the desired equality holds. On the other hand, if the right hand expression is finite, then the set $\phi(X_i, -T)$ vanishes as $T \to \infty$, so the limit expression on the left hand side is zero and the desired equality holds anyway.

This finishes the proof for piecewise constant l. The result follows by continuity for arbitrary l. □

We are now ready for the proof of Theorem 1.

Proof of Theorem 1 The equality (1) will be proved by separately deriving inequalities in the two opposite directions. First, the right hand side of (1) is proved to be at least as big as the left hand side. This is done by explicit construction of a control law in $\mathcal{U}$ from density functions ρ_i satisfying (2). The desired inequality then follows from Lemma 1.

Let γ^* be the value of the infimum in (1). Define $\rho_i \in C_0^1(\overline{X})$, $i = 1, \ldots, M$ such that (2) holds, $\rho_i > 0$ in X and

$$\sum_{i=1}^M \int_X l_i(x)\rho_i(x)dx < \gamma^* + \varepsilon$$

Define $u \in C^1(X)$ and $f \in C^1(X)$ according to

$$\rho(x) = \sum_i \rho_i(x) \qquad u_i(x) = \frac{\rho_i(x)}{\rho(x)}$$

$$f(x) = \sum_i u_i(x) f_i(x)$$

Then

$$\nabla \cdot (f\rho)(x) = \sum_i \nabla \cdot (f_i\rho_i)(x) > \psi(x) \qquad x \in X$$

and for every solution to the equation $\dot{x} = f(x)$

$$\begin{aligned}
0 \le \frac{\psi(x)}{\rho(x)} < \frac{\nabla \cdot (f\rho)}{\rho} &= \nabla \cdot f + \frac{\nabla \rho}{\rho} \cdot f \\
&= \nabla \cdot f + \nabla(\log \rho) \cdot f \\
0 < (\nabla \cdot f)(x(t)) &+ \frac{d}{dt} \log \rho(x(t))
\end{aligned}$$

By continuity of $\nabla \cdot f$ on the compact set $\overline{X}$, there is a constant C such that $\nabla \cdot f(x) \le C$ for $x \in X$. Hence

$$-C < \frac{d}{dt} \log \rho(x(t)) \qquad \rho(x(t)) > e^{-Ct}\rho(x(0))$$

This shows that the trajectory can not approach the boundary of X, where $\rho = 0$, but must stay in X for all $t \ge 0$. Hence by Lemma 1

$$\int_X \nabla \cdot (f\rho)V_u dx = \sum_{i=1}^M \int_X l_i(x)\rho_i(x)dx < \gamma^* + \varepsilon$$

The choice of $\varepsilon > 0$ was arbitrary, so

$$\int_X \psi(x)V^*(x)dx < \int_X \nabla \cdot (f\rho)V_u dx \le \gamma^*$$

To complete the proof, it remains to prove inequality in the opposite direction as well. For this purpose, define two subsets of $K = \mathbf{R} \times C(\overline{X})$:

$$\begin{aligned}
K_1 = \Bigg\{ \left(-\sum_{i=1}^M \int_X l_i\rho_i dx + \gamma^*, \sum_{i=1}^M \nabla \cdot (f_i\rho_i) - \psi \right) \Bigg| \\
\rho_i \in C_0^1(\overline{X}),\ \rho_i > 0 \text{ in } X \Bigg\}
\end{aligned}$$

$$K_2 = \{(z,h) \in K \mid z \ge 0,\ h(x) > 0 \text{ for } x \in X \setminus \Gamma\}$$

We will next prove the following statements:

I K_1 contains no interior point of K_2.

II There exists $k^* \in K^*, k^* \ne 0$ such that

$$\sup_{k_1 \in K_1} \langle k_1, k^* \rangle \le \inf_{k_2 \in K_2} \langle k_2, k^* \rangle, \tag{4}$$

where K^* is the dual space of K, i.e. $K^* = \mathbf{R} \times C(\overline{X})'$

III There exists a nonzero pair (a, ϕ) where $a \ge 0$ is a number and $\phi \ge 0$ is a measure of bounded variation on $\overline{X}$, vanishing inside Γ, such that $\langle \phi, \psi \rangle \ge a\gamma^*$ and for $i = 1, 2, \ldots, M$

$$al_i + \nabla\phi \cdot f_i \ge 0 \qquad \text{in } X \tag{5}$$

The derivative $\nabla\phi$ is interpreted in the sense of distributions.

The statement **I** is trivial once it is noted that $(z,h) \in K_2$ is an interior point if and only if $z > 0$ and $h(x) > 0$ for $x \in \overline{X}$.

The equivalence **I**$\Leftrightarrow$**II** holds because of the following separation property of convex sets [Luenberger, 1969; Rudin, 1991]: *Let K be a normed vector space and denote its dual K^*. Let K_1 and K_2 be convex sets in K such that K_2 has interior points and K_1 contains no interior point of K_2. Then there is a closed hyperplane separating K_1 and K_2; i.e., there is a $k^* \in K^*$, $k^* \neq 0$ such that*

$$\sup_{k_1 \in K_1} \langle k_1, k^* \rangle \leq \inf_{k_2 \in K_2} \langle k_2, k^* \rangle, \tag{6}$$

To show that **II**$\Leftrightarrow$**III**, let $k^* = (a, \phi) \in K^* = \mathbf{R} \times C(\overline{X})'$. The space $C(\overline{X})'$ is the set of measures of bounded variation and support in $\overline{X}$ [Dunford and Schwartz, 1958, page 262]. Expand the right hand side of (4) to

$$\inf_{k_2 \in K_2} \langle k_2, k^* \rangle = \inf_{(z,h) \in K_2} \{za + \langle \phi, h \rangle\} \tag{7}$$

The right hand side is equal to zero if and only if a and ϕ are both non-negative and $\phi = 0$ in Γ. Otherwise it is $-\infty$.

The left hand side of (4) can be expanded to

$$\begin{aligned}
&\sup_{k_1 \in K_1} \langle k^*, k_1 \rangle \\
&= \sup_{\rho_i} \Big\{ a\big(-\sum_i \int_X l_i \rho_i dx + \gamma^*\big) + \langle \phi, \sum_i \nabla \cdot (f_i \rho_i) - \psi \rangle \Big\} \\
&= \sup_{\rho_i} \Big\{ \sum_i \langle -a l_i - \nabla \phi \cdot f_i, \rho_i \rangle \Big\} + a\gamma^* - \langle \phi, \psi \rangle
\end{aligned}$$

The supremum is taken over $\rho_i \in C^1(\mathbf{R}^n)$ with support in $\overline{X}$ and $\rho_i > 0$ in X. The value is equal to $a\gamma^* - \langle \phi, \psi \rangle$ if and only if (5) holds, otherwise it is $+\infty$. The statement **II** is thus equivalent to **III**.

Notice that if the pair (a, ϕ) satisfies the conditions in **III** and $a = 0$, then $\langle \phi, \psi \rangle > 0$, so the conditions remain valid if a is replaced by a sufficiently small positive number. Hence a can always be assumed non-zero and be normalized to one.

Finally, let $u \in \mathcal{U}$ be arbitrary and define $f = \sum_i f_i u_i$ and $l = \sum_i l_i u_i$. Let $\rho \in C_0^1(X)$ be non-negative and such that $(\nabla \cdot (f\rho))(x) > \psi$ for $x \in \overline{X}$. Then

$$\begin{aligned}
&\int_X V_u(x)(\nabla \cdot (f\rho)(x) dx = \int_X l(x)\rho(x) dx \\
&= \langle l + \nabla \phi \cdot f, \rho \rangle + \langle \nabla \phi \cdot (f\rho) \rangle \\
&\geq \langle \phi, \nabla \cdot (f\rho) \rangle \\
&\geq \langle \phi, \psi \rangle \geq \gamma^*
\end{aligned}$$

The second equality uses $\langle \nabla \phi \cdot f, \rho \rangle = -\langle \phi, \nabla \cdot (f\rho) \rangle$ which holds by definition of distributional derivatives. This completes the proof. □

4. References

Dunford, N. and J. T. Schwartz (1958): *Linear Operators, Part I: General Threory.* Interscience Publishers, Inc.

Ford, L. K. and D. K. Fulkerson (1962): *Flows in Networks.* Princeton University Press, Princeton, New Jersey.

Hitchcock, F. L. (1941): "The distribution of a product from several sources to numerous localities." *J. Math. Phys.*, **20**, pp. 224–230.

Kantorovich, L. (1942): "On the transfer of masses." *Dokl. Akad. Nauk. SSSR*, **37**, pp. 227–229. In russian.

Luenberger, D. (1969): *Optimization by vector space methods.* Wiley, New York.

Rachev, S. and L. Rüschendorf (1998): *Mass Transporation Problems, Volume I: Theory.* Probability and its Applications. Springer.

Rantzer, A. (2001): "A dual to Lyapunov's stability theorem." *Systems & Control Letters*, **42:3**, pp. 161–168.

Rantzer, A. and F. Ceragioli (2001): "Smooth blending of nonlinear controllers using density functions." In *Proceedings of European Control Conference.*

Rantzer, A. and P. Parrilo (2000): "On convexity in stabilization of nonlinear systems." In *Proceedings of IEEE Conference of Decision and Control.*

Rudin, W. (1991): *Functional Analysis*, 2nd edition. McGraw-Hill.

Vinter, R. (1993): "Convex duality and nonlinear optimal control." *SIAM J. Control and Optimization*, **31:2**, pp. 518–538.

Young, L. C. (1969): *Lectures on the Calculus of Variations and Optimal Control Theory.* W. B. Saunders Company, Philadelphia, Pa.