WORKING PAPERS IN ECONOMICS

No. 05/07

S. D. FLÅM, H. TH. JONGEN AND
O. STEIN

SLOPES OF SHADOW PRICES AND LAGRANGE MULTIPLIERS

Department of Economics
UNIVERSITY OF BERGEN

Slopes of Shadow Prices and Lagrange Multipliers

S.D. Flåm* H.Th. Jongen# O. Stein‡

Abstract

Many economic models and optimization problems generate (endogenous) shadow prices - alias dual variables or Lagrange multipliers. Frequently the "slopes" of resulting price curves - that is, multiplier derivatives - are of great interest. These objects relate to the Jacobian of the optimality conditions. That particular matrix often has block structure. So, we derive explicit formulas for the inverse of such matrices and, as a consequence, for the multiplier derivatives.

Keywords: Sensitivity, optimal value function, shadow price, Karush-Kuhn-Tucker system, matrix inversion.

AMS classifications: primary 90C31, secondary 90C30, 90C90.

*Department of Economics, University of Bergen, Norway. Support from Ruhrgas is gratefully acknowledged.

#Department of Mathematics – C, RWTH Aachen University, Germany.

‡(Corresponding author) Department of Economics, University of Karlsruhe, Germany.
This author gratefully acknowledges financial support through a Heisenberg grant of the *Deutsche Forschungsgemeinschaft.*

1 Introduction

Comparative statics - alias sensitivity analysis - is crucial for economists (and engineers). Chief techniques include system studies or simulations, these fields providing most useful, often indispensable tools. If, however, the objective (or performance criterion) stems from optimization, then duality theory delivers derivative estimates with respect to parameter perturbations. On such occasions, first-order information is already embodied in Lagrange multipliers.

This feature is well-known, very convenient - and frequently fully satisfying. Some situations call though, for one step further down the road: They require second derivatives of the value (perturbation) function. To meet that request amounts to produce derivatives of Lagrange multipliers. Such derivatives are the main objects of this paper.

Our motivation stems from extremum problems of the following prototypical sort: Choose $x \in \mathbb{X}$ to

$$\text{optimize } f(x,t) \ \text{ subject to } h(x,t) = 0. \tag{1}$$

Here the objective function f is real-valued, $t \in \mathbb{T}$ is a parameter, and h maps $\mathbb{X} \times \mathbb{T}$ into $\mathbb{E}$. All spaces $\mathbb{X}$, $\mathbb{T}$, $\mathbb{E}$ are finite-dimensional Euclidean with inner products $\langle \cdot, \cdot \rangle$. For the applications we have in mind, f does not depend on t, and $h(x,t) = t - H(x)$ with a function H from $\mathbb{X}$ into $\mathbb{E} = \mathbb{T}$. We then choose $x \in \mathbb{X}$ to

$$\text{optimize } f(x) \ \text{ subject to } H(x) = t. \tag{2}$$

For the more general problem (1) consider the standard Lagrangian

$$L(x,t,\lambda) := f(x,t) + \langle \lambda, h(x,t) \rangle$$

and a Kuhn-Tucker (primal-dual) solution $t \mapsto (x(t), \lambda(t))$. Assuming differentiability, we mainly want to assess $\frac{d}{dt}\lambda(t)$. For this purpose one could first introduce the optimal value function

$$v(t) := opt_x \left\{ f(x,t) : h(x,t) = 0 \right\};$$

second, argue that in the specially structured case (2) one has $\lambda(t) = v'(t)$ - and finally, identify $\lambda'(t) = v''(t)$. This plan presumes however, that $v(\cdot)$ be twice differentiable. In fact, quite often v isn't even differentiable. So, the said plan may encounter formidable hurdles. To mitigate these we posit that f, h be at least C^2. Further, assume the constraint qualification that $D_x h(x,t)$ has full row rank at any optimal x. Plainly, problem (1) is fairly tractable, featuring neither restricted decision sets nor inequality constraints.[1] Nonetheless, its format is frequent and important enough to merit separate treatment.

For motivation Section 2 brings out four examples, all of micro-economic or game theoretic sort. In these, as in manifold other instances, the Jacobian of the optimality conditions comes as a block-structured matrix. Section 3 therefore prepares the ground by inverting a suitable class of such matrices. Section 4 applies that inversion to estimate parameter sensitivity of primal-dual solutions in smooth, equality-constrained optimization, phrased in the forms (1) and (2). We illustrate the results with an example.

While the findings in this paper may partly be seen as consequences of more general results from the literature, our main concern is to connect these second order sensitivity results with important applications from economics and, thus, provide insights for audiences from both areas.

2 Motivation

Example 1: Risk aversion in the small [21]**.** Consider an economic agent who maximizes his utility $u(x)$ subject to $\mathcal{A}x = t + \Delta t$. The matrix $\mathcal{A}$ has merely one row, x is a column vector of appropriate size, t is a real constant, and Δt is a random variable, called a *risk*, with expectation $E\Delta t = 0$. The objective $u(\cdot)$ is concave, whence so is the associated reduced function

$$U(t + \Delta t) := \sup \{u(x) : \mathcal{A}x = t + \Delta t\} \ ,$$

emerging ex post, after Δt has been unveiled. Since, by Jensen's inequality, $EU(t + \Delta t) \leq U(t)$, the agent displays risk aversion ex ante. He is then willing to pay a premium for avoiding uncertainty. Define that premium Π by $EU(t + \Delta t) = U(t - \Pi)$. To estimate Π, assume differentiability, and develop both sides of the last equation. Doing so yields

$$E\left\{U(t) + U'(t)\Delta t + U''(t)\Delta t^2/2\right\} \approx EU(t+\Delta t) = U(t-\Pi) \approx U(t) - U'(t)\Pi$$

and thereby

$$\Pi \approx -\frac{U''(t)}{U'(t)} var(\Delta t)/2.$$

[1]Also, since f is finite-valued, no implicit constraints are embodied by means of infinite penalties.

The quotient $-U''(t)/U'(t)$ is called the *Arrow-Pratt measure of absolute risk aversion.*[2] Under appropriate conditions there exists a Lagrange multiplier λ, satisfying $U'(t) = \lambda$. Suppose the mapping $t \mapsto \lambda(t)$ so defined be differentiable. Then

$$\Pi \approx -\frac{\lambda'(t)}{\lambda(t)} var(\Delta t)/2. \tag{3}$$

Example 2: Production games [6]. Suppose individual $i \in I$, $2 \leq |I| < +\infty$, faces a "private production task" t_i, construed as an obligation to supply a resource bundle (vector) t_i in a finite-dimensional Euclidean space $\mathbb{E}$, equipped with inner product $\langle \cdot, \cdot \rangle$. If supplying $x_i \in \mathbb{E}$, he incurs cost $C_i(x_i) \in \mathbb{R} \cup \{+\infty\}$. Members of any *coalition* $S \subseteq I$ could pool their endowments, coordinate their efforts, and thereby generate aggregate cost

$$C_S(t_S) := \inf \left\{ \sum_{i \in S} C_i(x_i) : \sum_{i \in S} x_i = t_S \right\}, \tag{4}$$

with $t_S := \sum_{i \in S} t_i$. Construction (4), being crucial in nonlinear analysis, is commonly called an *inf-convolution;* see [19]. A cost-sharing scheme $(c_i) \in \mathbb{R}^I$ resides in the *core* iff

$$\begin{array}{lll} \textit{Pareto efficient}: & \sum_{i \in I} c_i = C_I(t_I), & \text{and} \\ \textit{socially stable}: & \sum_{i \in S} c_i \leq C_S(t_S) & \text{for all } S \subset I. \end{array}$$

Suppose $\lambda \in \mathbb{E}$ is a Lagrange multiplier that relaxes the coupling constraint in (4) when $S = I$. More precisely, suppose

$$\inf_x \sum_{i \in I} \{C_i(x_i) + \langle \lambda, t_i - x_i \rangle\} \geq C_I(t_I).$$

Let $C_i^*(\lambda) := \sup_{x_i} \{\langle \lambda, x_i \rangle - C_i(x_i)\}$ denote the *Fenchel conjugate* of C_i. The profile

$$i \mapsto c_i := \langle \lambda, t_i \rangle - C_i^*(\lambda) \tag{5}$$

then belongs to the core [6]. Existence of a Lagrange multiplier λ is ensured if $C_I(\cdot)$ is finite-valued in a neighborhood around t_I and convex.

For a monopolistic setting of this story, suppose the agents are parallel branches of an integrated concern, gaining aggregate revenue $R(t)$ when

[2]This is a local measure. It emerged first in studies by de Finetti (1952), Arrow (1963) and Pratt (1964). The agent at hand would regard a small, symmetric risk Δt around the specified level t as equivalent to retaining t less $-\frac{1}{2} var(\Delta t)U''(t)/U'(t)$. The inverse quantity $\mathcal{T}(t) := -U'(t)/U''(t)$ is called the *risk tolerance* [20].

putting out total production volume t. To verify its second order optimality - or to test for possible risk aversion - the said concern would look at $R''(t) - C_I''(t)$. Assume C_I has a second Fréchet-derivative in a neighborhood of t which is continuous and non-singular at that point. Then, if C_I is convex, by a result of Crouzeix [3],

$$C_I''(t) = \left[C_I^{*''}(t^*)\right]^{-1}$$

where $t^* = \lambda = C_I'(t)$ and $C_I^*(t^*) = \sum_{i \in I} C_i^*(t^*)$. For generalization, see [22, Theorem 13.21]. It follows, under quite similar assumptions on the C_i, that

$$C_I^{*\prime\prime}(t^*) = \sum_{i \in I} C_i^{*\prime\prime}(t^*) = \sum_{i \in I} \left[C_i''(x_i)\right]^{-1}$$

where $x_i, i \in I$, is the supposedly unique, feasible profile that yield total cost $C_I(t)$. The upshot is that

$$\lambda'(t) = C_I''(t) = \left[\sum_{i \in I} \left[C_i''(x_i)\right]^{-1}\right]^{-1}. \tag{6}$$

For interpretation and analogy regard $C_i''(x_i)$ as the "resistance" in branch i, its inverse being the corresponding "conductance" there. Formula (6) then points to electrical engineering, saying that the conductance of a parallel circuit equals the sum of conductances [5].

For a quite opposite, perfectly competitive setting, let $i \in I$ be independent firms, each acting as a price-taking supplier in common product markets. These markets clear at price $p = \lambda = C'(t_I)$, and marginal costs are then equal across the industry: $p = C_i'(x_i)$ for each smooth-cost firm i having optimal choice x_i interior to the domain where C_i is finite-valued. Let $\mathbb{E} = \mathbb{R}^G$ for a finite set G of goods. Fixing any two goods $g, \bar{g} \in G$ the *demand elasticity* of the first good with respect to the price of the second is defined by

$$\varepsilon_{g\bar{g}} := \lim_{\Delta p_{\bar{g}} \to 0} \frac{\Delta t_g / t_g}{\Delta p_{\bar{g}} / p_{\bar{g}}} = \frac{p_{\bar{g}}}{t_g} \left[\frac{\partial p_{\bar{g}}}{\partial t_g}\right]^{-1} = \frac{\lambda_{\bar{g}}}{t_g} \left[\frac{\partial \lambda_{\bar{g}}}{\partial t_g}\right]^{-1}.$$

Important *market games* [23] obtain by rather putting profit (instead of cost) at center stage. Specifically, if agent $i \in I$ owns resources $t_i \in \mathbb{E}$, and enjoys payoff $\pi_i : \mathbb{E} \to \mathbb{R} \cup \{-\infty\}$, the characteristic form TU-game

$$S \subseteq I \mapsto \pi_S(t_S) := \sup \left\{ \sum_{i \in S} \pi_i(x_i) : \sum_{i \in S} x_i = t_S \right\}, \tag{7}$$

has for each $\lambda \in \mathbb{E}$, satisfying

$$\sup_x \sum_{i \in I} \{\pi_i(x_i) + \langle \lambda, t_i - x_i \rangle\} \leq \pi_I(t_I),$$

a core solution $(c_i) \in \mathbb{R}^I$ defined by

$$c_i := \sup_{x_i} \{\pi_i(x_i) + \langle \lambda, t_i - x_i \rangle\} = \langle \lambda, t_i \rangle + (-\pi_i)^*(-\lambda).$$

That is, $\sum_{i \in S} c_i \geq \pi_S(t_S)$ for all $S \subset I$ with equality when $S = I$. Such market games are vehicles in studies of welfare gains from trading natural resources, be the latter fish quotas or pollution permits [8].

Example 3: Mutual insurance [12], [13]. Next relate the preceding two examples as follows. Let $\pi_i(x) := E\Pi_i(\omega, x_i(\omega))$ denote the expected payoff of agent i when enjoying a state-contingent wealth profile $\omega \in \Omega \mapsto x_i(\omega) \in \mathbb{R}$. Here Ω is a finite state space, equipped with probability measure $\omega \mapsto \Pr(\omega) > 0$. Correspondingly, let $\mathbb{E} := \mathbb{R}^\Omega$ have probabilistic inner product $\langle e, e' \rangle := \sum_{\omega \in \Omega} e(\omega)e'(\omega) \Pr(\omega)$. Agent i now owns a *risk* $t_i \in \mathbb{E}$. In that optic a Lagrange multiplier $\lambda \in \mathbb{E}$ has twin properties: *After* state ω has been unveiled it holds for

$$c_i(\omega) := \sup_{x_i} \{\Pi_i(\omega, x_i) + \lambda(\omega) [t_i(\omega) - x_i]\}$$

and

$$\Pi_S(\omega, t_S(\omega)) := \sup \left\{ \sum_{i \in S} \Pi_i(\omega, x_i) : \sum_{i \in S} x_i = t_S(\omega) \right\} \tag{8}$$

that $\sum_{i \in S} c_i(\omega) \geq \Pi_S(\omega, t_S(\omega))$ for all $S \subset I$ with equality for $S = I$. Thus the contingent payment profile $[c_i(\omega)] \in \mathbb{R}^I$ belongs to the core of the ex post, second-stage game defined by characteristic function (8).

Similarly, *before* ω is known the expected payments Ec_i belongs to the ex ante game with characteristic function defined by (7). Together the scheme (c_i, Ec_i) might justly be called a mutual insurance treaty. Posit that each Π_i be state independent, this meaning that $\Pi_i(\omega, \cdot) = \Pi_i(\cdot)$. Further, let $t_i = \bar{t}_i + \Delta t_i$ with $\bar{t}_i$ known, Δt_i random, and $E\Delta t_i = 0$. Then the *risk tolerance* of the mutual company at $\bar{t}_I := \sum_{i \in I} \bar{t}_i$ equals $-\lambda/\lambda' =$

$$\mathcal{T}_I := -\frac{\Pi_I'}{\Pi_I''} = -\sum_{i \in I} \frac{\Pi_i'}{\Pi_i''} = \sum_{i \in I} \mathcal{T}_i,$$

in compliance with formula (6).

Example 4: Playing the market [9]**.** Instead of agents $i \in I$ all being part of one corporation (or mutual), suppose now that these parties compete in the following manner. At a first stage firm i independently commits to supply the commodity vector $t_i \in \mathbb{E}$. By doing so it gains gross revenue $R_i(\mathbf{t})$ in the market, $\mathbf{t} := (t_i) \in \mathbb{E}^I$ denoting the profile of commitments. Production cost must be covered though. So, for the sake of efficiency and fair sharing, after $\mathbf{t}$ has already been committed, firms collaborate and split costs as described by (5). Consequently, the final payoff to firm i equals

$$\pi_i(\mathbf{t}) := R_i(\mathbf{t}) - \langle \lambda, t_i \rangle + C_i^*(\lambda),$$

$\lambda \in \mathbb{E}$ being a Lagrange multiplier associated to the problem (4) when $S = I$. Let $\mathbf{t}_{-i}$ be short notation for $(t_j)_{j \neq i}$ and declare the profile $\mathbf{t}$ a *Nash equilibrium* if for each i

$$t_i \text{ maximizes } R_i(\mathbf{t}_{-i}, \cdot) - \langle \lambda, \cdot \rangle + C_i^*(\lambda).$$

It is tacitly understood here that λ depends on $t_I = \sum_{i \in I} t_i$. We posit that each party fully knows that feature. Assuming differentiability, the first order optimality conditions for equilibrium read: for each $i \in I$,

$$\begin{aligned} \frac{\partial}{\partial t_i} R_i(\mathbf{t}_{-i}, t_i) &= \lambda + \lambda'(t_I) t_i + C_i^{*\prime}(\lambda) \lambda'(t_I) \\ &= \lambda + \lambda'(t_I) t_i + x_i \lambda'(t_I), \end{aligned}$$

x_i being the supposedly unique choice in (4) when $S = I$. For λ' one may apply formula (6). Admittedly, the issues concerning existence and uniqueness of such Nash equilibrium are intricate. For discussion of these issues see [8], [10].

In the above examples problem (1) assumes the simpler form (2) for which λ is commonly called a *shadow price*. Since our results about derivatives of optimal value functions and Lagrange multipliers can be obtained for the more general problem (1) without much additional effort, we will concentrate on this setting and give the simplified formulas for instance (2) in the end.

As will be shown in Section 4 in more detail, the KKT conditions of problem (1), namely:

$$\begin{array}{lll} D_x f(x,t) & + \left\langle \lambda, D_x^\top h(x,t) \right\rangle & = 0 \\ h(x,t) & & = 0 \end{array}$$

generate a Jacobian with block form

$$\begin{pmatrix} A & B \\ B^\top & 0 \end{pmatrix}$$

featuring $A = D_x^2 L(x,t,\lambda)$ and $B^\top = D_x h(x,t)$. This simple observation leads us to inquire next about inversion of such matrices.

3 The inverse of a structured block matrix

Definition 3.1 *For an $(n,n)-$matrix A and an $(n,k)-$matrix B with $k < n$ the restriction of A to the kernel of $B^\top$ is defined as*

$$A|_{\ker(B^\top)} \; = \; V^\top A V \; ,$$

where V denotes any matrix whose columns form a basis of $\ker(B^\top)$.

Remark 3.2 *In the following we will only be interested in properties of $A|_{\ker(B^\top)}$ which do not depend on the actual choice of V.*

A proof for the well known part a) of the following theorem can be found in [17]. Part b) was first shown in [15] under more general assumptions, requiring an elaborate proof technique. In fact, there the Moore-Penrose inverse of Q is given for the case that B does not possess full rank. In contrast, here we give an elementary proof for a problem structure which is adequate for the applications we have in mind.

Theorem 3.3 *Let A be an $(n,n)-$matrix, let B be an $(n,k)-$matrix with $k < n$, and define the block matrix*

$$Q \; = \; \left(\begin{array}{c|c} A & B \\ \hline B^\top & 0 \end{array}\right) \; .$$

a) Q is nonsingular if and only if $\operatorname{rank}(B) = k$ *and $A|_{\ker(B^\top)}$ is nonsingular.*

b) Let Q be nonsingular and let the columns of V form a basis of $\ker(B^\top)$. *Then, with*

$$W \; = \; V(V^\top A V)^{-1} V^\top \quad \text{and} \quad M \; = \; (B^\top B)^{-1} B^\top \; ,$$

the inverse of Q is given by

$$Q^{-1} = \left(\begin{array}{c|c} W & (I-WA)\,M^\top \\ \hline M\,(I-AW) & M\,(AWA-A)\,M^\top \end{array}\right).$$

Proof of part b). By part a), we have $\mathrm{rank}(B) = k$ so that M is well defined and V is an $(n, n-k)-$matrix with $\mathrm{rank}(V) = n-k$ and $V^\top B = 0$. Also by part a), the matrix $V^\top AV$ is nonsingular, so that W is well defined, too. Now consider the equation

$$\left(\begin{array}{c|c} A & B \\ \hline B^\top & 0 \end{array}\right)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$$

or, equivalently, the system

$$Ax + By = c \tag{9}$$
$$B^\top x = d \tag{10}$$

A basis for the homogeneous part of (10) is given by the columns of V, and it is easily verified that $B(B^\top B)^{-1}d$ is a particular solution. Thus the solutions of (10) are given as

$$x = V\xi + B(B^\top B)^{-1}d \tag{11}$$

with $\xi \in \mathbb{R}^{n-k}$. Plugging (11) into (9) and multiplying by $V^\top$ from the left yields an equation which can be solved for ξ, and (11) yields

$$x = W\,c + (I - WA)\,M^\top d\,. \tag{12}$$

After plugging (12) into (9) and multiplying by $B^\top$ from the left one obtains an equation that can be solved for y, so that finally one has

$$\begin{pmatrix} x \\ y \end{pmatrix} = Q^{-1}\begin{pmatrix} c \\ d \end{pmatrix}$$

with the claimed matrix Q^{-1}. •

Remark 3.4 *In Theorem 3.3b), M is the Moore-Penrose inverse of B.*

Remark 3.5 *Given A and B, in Theorem 3.3b) the matrices W and, thus, Q^{-1} do not depend on the actual choice of V. Note that we have $WAW = W$ and $(AWA)|_{\ker(B^\top)} = A|_{\ker(B^\top)}$, so that A is a generalized inverse (shortly: g-inverse [11]) of W, and W is a "restricted g-inverse" of A.*

Remark 3.6 *In the case $k = n$ the matrix V from Definition 3.1 cannot be defined (formally, $A_{\ker(B^\top)}$ is then a "$(0,0)-$matrix"). The corresponding result about nonsingularity and the inverse of Q is, however, easily derived: Let A and B be $(n,n)-$matrices. Then the matrix*

$$Q \;=\; \left(\begin{array}{c|c} A & B \\ \hline B^\top & 0 \end{array}\right)$$

is non-singular if and only if B is nonsingular. In the latter case, the inverse of Q is given by

$$Q^{-1} \;=\; \left(\begin{array}{c|c} 0 & (B^\top)^{-1} \\ \hline B^{-1} & -B^{-1}A(B^\top)^{-1} \end{array}\right) \;.$$

Remark 3.7 *Results about the inertia of a matrix structured like Q in Theorem 3.3 can be found in [16].*

4 Derivatives of Lagrange multipliers

Returning to problem (1), in this section we consider the parametric optimization problem

$$P(t): \qquad \min_x \; f(x,t) \quad \text{subject to} \quad h(x,t) \;=\; 0$$

with $x \in \mathbb{X} := \mathbb{R}^n$, $t \in \mathbb{T} := \mathbb{R}^r$, and functions $f \in C^2(\mathbb{R}^n \times \mathbb{R}^r, \mathbb{R})$ and $h \in C^2(\mathbb{R}^n \times \mathbb{R}^r, \mathbb{E})$ with $\mathbb{E} = \mathbb{R}^k$ and $k < n$. Let $\bar{x}$ be a nondegenerate critical point of $P(\bar{t})$, that is, there exists some $\bar{\lambda} \in \mathbb{R}^k$ with

$$D_x f(\bar{x},\bar{t}) + \bar{\lambda}^\top D_x h(\bar{x},\bar{t}) \;=\; 0 \;,$$

the matrix $D_x h(\bar{x},\bar{t})$ has full rank k, and the restricted Hessian $D_x^2 L(\bar{x},\bar{t},\bar{\lambda})|_{\ker(D_x h(\bar{x},\bar{t}))}$ is nonsingular, where

$$L(x,t,\lambda) \;=\; f(x,t) + \lambda^\top h(x,t) \;.$$

Putting $A = D_x^2 L(\bar{x},\bar{t},\bar{\lambda})$ and $B = D_x^\top h(\bar{x},\bar{t})$, under these assumptions Theorem 3.3a) implies the nonsingularity of the matrix

$$Q \;=\; \left(\begin{array}{c|c} A & B \\ \hline B^\top & 0 \end{array}\right) \;.$$

We emphasize that nondegeneracy of a critical point is a weak assumption. For example, when the defining functions are in general position, for parameterfree problems all critical points are nondegenerate, and for oneparametric problems almost all critical points are nondegenerate ([14]). Moreover, if f is strictly convex in x with $D_x^2 f(x,t)$ positive definite for all x and t and if, in addition, h is linear in x with $D_x h(x,t) = A(t)$, then the only critical point of $P(t)$ (its global minimizer) is nondegenerate whenever $A(t)$ has full rank.

As Q is the Jacobian with respect to (x, λ) of the system

$$\begin{aligned} D_x^\top f(x,t) + D_x^\top h(x,t)\,\lambda &= 0 \\ h(x,t) &= 0 \end{aligned}$$

at $(\bar{x}, \bar{t}, \bar{\lambda})$, the implicit function theorem and a moment of reflection show that for t close to $\bar{t}$ there exists a locally unique nondegenerate critical point $x(t)$ of $P(t)$ with multiplier $\lambda(t)$. In particular, the functions $x(t)$ and $\lambda(t)$ satisfy the equations

$$D_x^\top f(x(t),t) + D_x^\top h(x(t),t)\,\lambda(t) = 0 \tag{13}$$

$$h(x(t),t) = 0 \tag{14}$$

for all t in a neighborhood of $\bar{t}$.

Assuming that $\bar{x}$ is even a nondegenerate local minimizer of $P(\bar{t})$, it is not hard to see that $x(t)$ is a local minimizer of $P(t)$ for t close to $\bar{t}$. Hence the (local) optimal value function of $P(t)$ is

$$v(t) = f(x(t),t) \ .$$

In order to calculate the derivative of v observe that by (14) we may also write

$$v(t) = f(x(t),t) + \lambda(t)^\top h(x(t),t)$$

which yields

$$\begin{aligned} v'(t) &= D_x L(x(t),t,\lambda(t))\,x'(t) + \lambda'(t)^\top h(x(t),t) + D_t L(x(t),t,\lambda(t)) \\ &= D_t L(x(t),t,\lambda(t)) \ , \end{aligned} \tag{15}$$

where we used (13) and (14). For the second derivative of v at $\bar{t}$ we obtain by differentiation of (15)

$$v''(\bar{t}) = D_t^2 L(\bar{x},\bar{t},\bar{\lambda}) + \begin{pmatrix} D_t D_x^\top L(\bar{x},\bar{t},\bar{\lambda}) \\ D_t h(\bar{x},\bar{t}) \end{pmatrix}^\top \begin{pmatrix} x'(\bar{t}) \\ \lambda'(\bar{t}) \end{pmatrix} .$$

As differentiation of (13) and (14) yields

$$Q \begin{pmatrix} x'(\bar{t}) \\ \lambda'(\bar{t}) \end{pmatrix} + \begin{pmatrix} D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) \\ D_t h(\bar{x}, \bar{t}) \end{pmatrix} \;=\; 0 \;,$$

we arrive at the formula

$$v''(\bar{t}) \;=\; D_t^2 L(\bar{x}, \bar{t}, \bar{\lambda}) \;-\begin{pmatrix} D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) \\ D_t h(\bar{x}, \bar{t}) \end{pmatrix}^\top Q^{-1} \begin{pmatrix} D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) \\ D_t h(\bar{x}, \bar{t}) \end{pmatrix} , \tag{16}$$

where a so-called shift term is subtracted from the Hessian of L with respect to t (cf. [17] for details).

With a matrix V whose columns form a basis of $\ker(B^\top) = \ker(D_x h(\bar{x}, \bar{t}))$ we can now evoke Theorem 3.3 to state explicit formulas for these derivatives:

$$\begin{aligned} \begin{pmatrix} x'(\bar{t}) \\ \lambda'(\bar{t}) \end{pmatrix} \;&=\; -Q^{-1} \begin{pmatrix} D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) \\ D_t h(\bar{x}, \bar{t}) \end{pmatrix} \\ &=\; \begin{pmatrix} -W\, D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) - (I - WA)\, M^\top\, D_t h(\bar{x}, \bar{t}) \\ -M\,(I - AW)\, D_t D_x^\top L(\bar{x}, \bar{t}, \bar{\lambda}) - M\,(AWA - A)\, M^\top\, D_t h(\bar{x}, \bar{t}) \end{pmatrix} \end{aligned}$$

with the notation from Theorem 3.3. More explicitly, the derivative of the Lagrange multiplier is

$$\begin{aligned} \lambda'(\bar{t}) \;&=\; \big(D_x\bar{h} D_x^\top \bar{h}\big)^{-1} D_x\bar{h}\, \big(D_x^2\bar{L} V (V^\top D_x^2 \bar{L} V)^{-1} V^\top - I\big)\, D_t D_x^\top \bar{L} \qquad (17) \\ &\quad + \big(D_x\bar{h} D_x^\top \bar{h}\big)^{-1} D_x\bar{h}\, \big(D_x^2\bar{L} - D_x^2\bar{L} V \big(V^\top D_x^2 \bar{L} V\big)^{-1} V^\top D_x^2 \bar{L}\big)\, D_x^\top \bar{h}\, \big(D_x\bar{h} D_x^\top \bar{h}\big)^{-1}\, D_t\bar{h} \end{aligned}$$

where $D_x\bar{h}$ stands for $D_x h(\bar{x}, \bar{t})$, etc.

Finally, we consider the special case of problem (2) with f independent of t and $h(x,t) = t - H(x)$. It is easily seen that (15) now yields the well-known result

$$v'(t) \;=\; \lambda(t) \;. \tag{18}$$

Moreover, (16) reduces to

$$\begin{aligned} v''(\bar{t}) \;&=\; 0 - \begin{pmatrix} 0 \\ I \end{pmatrix}^\top Q^{-1} \begin{pmatrix} 0 \\ I \end{pmatrix} \qquad (19) \\ &=\; \big(D_x\bar{H} D_x^\top \bar{H}\big)^{-1} D_x\bar{H}\, \big(D_x^2\bar{L} - D_x^2\bar{L} V \big(V^\top D_x^2 \bar{L} V\big)^{-1} V^\top D_x^2 \bar{L}\big)\, D_x^\top \bar{H}\, \big(D_x\bar{H} D_x^\top \bar{H}\big)^{-1} . \end{aligned}$$

Clearly, a combination of (17) and (18) would yield the same result for $v''(\bar{t}) = \lambda'(\bar{t})$, as we proposed in the introduction.

We illustrate the consequence of this formula for Example 1 from Section 2.

Example 1, continued: For the approximation of the premium Π we consider the above problem $P(t)$ with $f(x,t) = -u(x)$, $h(x,t) = t - H(x)$, and $H(x) = \mathcal{A}x$. The assumptions that u is strictly concave with $D^2u(x)$ negative definite for all x, and that the vector $\mathcal{A}$ does not vanish, are usually satisfied in applications. Then for all t each critical point of $P(t)$ is nondegenerate. The corresponding Lagrange function is $L(x,t,\lambda) = -u(x) + \lambda(t - \mathcal{A}x)$.

Let $\bar{x}$ be a nondegenerate critical point of $P(\bar{t})$. It is not hard to see that the multiplier satisfies

$$\lambda(\bar{t}) \;=\; -\frac{D\bar{u}\mathcal{A}^\top}{\mathcal{A}\mathcal{A}^\top}\,. \tag{20}$$

Moreover, we have $D_x^2 L(x,t,\lambda) = -D^2u(x)$ so that, with a basis matrix V of $\ker(\mathcal{A})$, formula (19) yields

$$\lambda'(\bar{t}) \;=\; \frac{1}{(\mathcal{A}\mathcal{A}^\top)^2}\mathcal{A}\left(-D^2\bar{u} + D^2\bar{u}V(V^\top D^2\bar{u}V)^{-1}V^\top D^2\bar{u}\right)\mathcal{A}^\top. \tag{21}$$

Plugging (20) and (21) into (3) leads to

$$\Pi \approx \frac{\mathcal{A}\left(D^2\bar{u}V(V^\top D^2\bar{u}V)^{-1}V^\top D^2\bar{u} - D^2\bar{u}\right)\mathcal{A}^\top}{\mathcal{A}\mathcal{A}^\top \cdot D\bar{u}\mathcal{A}^\top} var(\Delta t)/2.$$

5 Final remarks

The approach to use the implicit function theorem in parametric optimization goes back to [7]. We emphasize that it can also be carried out for parametric optimization problems with finitely many inequality constraints, when strict complementarity slackness is added to the nondegeneracy assumptions at a critical point. Our results about the inverse of the Jacobian and the multiplier derivatives then remain unchanged if the set of equality constraints is extended by the active inequality constraints.

Instead of differentiability of primal-dual solutions one might contend with Lipschitz behavior. On that issue, see [2] for additive and linear perturbations of convex problems, and [18] for general perturbations of nonconvex problems.

Acknowledgements

We would like to thank the anonymous referees for their precise and substantial remarks which lead to a significantly improved version of this article.

References

[1] K. J. Arrow, Liquidity preference, Lecture IV in Lecture Notes for Economics 285, *The Economics of Uncertainty* 33-53, Stanford University (1963).

[2] J-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, *Mathematics of Operations Research* 9, 87-111 (1984).

[3] J. P. Crouzeix, A relationship between the second derivatives of a convex function and of its conjugate, *Mathematical Programming* 13, 364-365 (1977).

[4] B. de Finetti, Sulla preferibilita, *Giornale degli Economisti e Annali di Economia* 11, 685-709 (1952).

[5] R. C. Dorf and J. A. Svoboda, *Introduction to Electric Circuits*, J. Wiley & Sons, New York (1999).

[6] I. V. Evstigneev and S. D. Flåm, Sharing nonconvex cost, *J. Global Optimization* 20, 257-271 (2001).

[7] A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques,* Wiley, New York, 1968.

[8] S. D. Flåm and O. Godal, Greenhouse Gases, Quota Exchange and Oligopolistic Competition, in C. Carraro and V. Fragnelli (Eds.), *Game Practice and the Environment*, Edward Elgar, Cheltenham, UK (2004).

[9] S. D. Flåm and A. Jourani, Strategic behavior and partial cost sharing, *Games and Economic Behavior* 43, 44-56 (2003).

[10] J. J. Gabszewicz, *Strategic Multilateral Exchange*, Edward Elgar, Cheltenham, UK (2002).

[11] D. A. Harville, *Matrix Algebra from a Statistician's Perspective*, Springer, New York (1999).

[12] D. Henriet and J-C. Rochet, Some reflexions on insurance pricing, *European Economic Review* 31, 863-885 (1987).

[13] D. Henriet and J-C. Rochet, *Microeconomie de l'assurance*, Economica, Paris (1991).

[14] H. Th. Jongen, P. Jonker, F. Twilt, One-parameter families of optimization problems: equality constraints, *Journal of Optimization Theory and Applications* 48, 141-161 (1986).

[15] H. Th. Jongen, T. Möbert and K. Tammer, On iterated optimization in nonconvex optimization, *Math. of Op. Research* 11, 679-691 (1986).

[16] H. Th. Jongen, T. Möbert, J. Rückmann and K. Tammer, On inertia and Schur complement in optimization, *Lin. Alg. Applications* 95, 97-109 (1987).

[17] H. Th. Jongen, K. Meer and E. Triesch, *Optimization Theory*, Kluwer, Boston (2004).

[18] D. Klatte and B. Kummer, *Nonsmooth equations in optimization*, Kluwer, Boston (2002).

[19] P-J. Laurent, *Approximation et optimisation*, Hermann, Paris (1972).

[20] M. Magill and M. Quinzii, *Theory of Incomplete Markets*, The MIT Press, Cambridge, Mass. (1996).

[21] J. W. Pratt, Risk aversion in the small and the large, *Econometrica* 32, 122-136 (1964).

[22] R. T. Rockafellar and J.-B. Wets, *Variational Analysis*, Springer, Berlin (1998).

[23] L. S. Shapley and M. Shubik, On market games, *J. Economic Theory* 1, 9-25 (1969).

Department of Economics
University of Bergen
Fosswinckels gate 6
N-5007 Bergen, Norway
Phone: +47 55 58 92 00
Telefax: +47 55 58 92 10
http://www.svf.uib.no/econ