Ricerche di Matematica 55: 1–12 (2006)
DOI: 10.1007/s11587-006-0001-2

Ivar Ekeland

A duality theory for some non-convex functions of matrices

Received: February 8, 2006 / Accepted: March 8, 2006

Abstract We study a special class of non-convex functions which appear in nonlinear elasticity, and we prove that they have a well-defined Legendre transform. Several examples are given, and an application to a nonlinear eigenvalue problem.

Keywords Duality · Legendre transform · Nonlinear Elasticity

Mathematics Subject Classification (2000) 05C38 · 15A15 · 05A15 · 15A18

1 Introduction

We want to define a Legendre transform $F^L(y)$ for functions $F(x)$, where x is an $N \times K$ matrix, and F involves the various cofactors of x. Note that F then has to be strongly nonlinear, and nonconvex. The simplest case is when $N = K$ and F is a function of the determinant only: $F(x) = \Phi(\det x)$. We show that the Legendre transform of

$$F(x) = \frac{N}{p} |\det x|^{p/N}$$

is

$$F^L(y) = \frac{N}{q} |\det y|^{q/N}$$

with $1/p + 1/q = 1$, thereby generalizing the classical duality between L^p spaces. The next simplest is when F is a function of the $(N-1)$-cofactors of x: in the case when $N = K = 3$, we give conditions under which F has a well-defined Legendre

Communicated by R. De Arcangelis

I. Ekeland
Canada Research Chair in Mathematical Economics
Pacific Institute of Mathematical Sciences, 1933 West Mall, University of British Columbia, Vancouver BC, Canada V6T 1Z2
E-mail: ekeland@math.ubc.ca

transform. This covers for instance the area functional, defined over 2×3 matrices $x = \left(x_j^i\right)$ by:

$$F(x) = \left[\left(x_1^1 x_2^2 - x_1^2 x_2^1\right)^2 + \left(x_1^1 x_3^2 - x_1^2 x_3^1\right)^2 + \left(x_2^1 x_3^2 - x_2^2 x_3^1\right)^2\right]^{1/2}$$

which turns out to be its own Legendre transform (in other words, it is self-dual).

Note that with our definition, the Legendre transform F^L of F satisfies the usual duality relations:

$$\left(F^L\right)^L = F$$

$$\left(F'\right)^{-1} = \left(F^L\right)'.$$

The paper is organized as follows. First we define precisely what we mean by a Legendre transform. Then we study two polar cases. In the first one F is a function of the determinant only, and in the second F depends only on the 2×2 cofactors of a 3×3 matrix. We give several examples, and we conclude by giving an application to a nonlinear eigenvalue problem.

2 The Legendre transform

Let X be a finite-dimensional vector space, and Y its dual, the duality pairing being denoted by $< x, y >$. Let $F : X \to R$ be a C^1 function and $F'(x) \in Y$ its derivative at x. The classical formula of Legendre associates with every $y \in Y$ a set $\Gamma_L(y) \subset R$ defined as follows:

$$\Gamma_L(y) = \left\{< x, y > - F(x) \mid y = F'(x)\right\}. \tag{1}$$

Usually, the right-hand contains several points, so that Γ_L is a multi-valued map from Y to R. We refer to [1] for a study of this map. For certain classes of functions F, however, the right-hand side is a singleton, so that formula (1) defines a function on Y, which is then called the Legendre transform of F.

Definition 1 Consider a map $F : \Omega \to R$, where Ω is a submanifold of X, and set $\Sigma = F'(\Omega)$. We shall say that a C^1 function $G : \Sigma \to R$ is the Legendre transform of F if Σ is a submanifold of Y and:

$$\left[x \in \Omega,\ y = F'(x)\right] \Longrightarrow < x, y > - F(x) = G(y).$$

Of course, if F' is one-to-one, this formula becomes:

$$G(y) = < \left(F'\right)^{-1}(y), y > - F\left(\left(F'\right)^{-1}(y)\right)\ \forall y \in \left(F'\right)^{-1}(\Omega).$$

We shall denote the Legendre transform of F by F^L, so that $F^L = G$ in the above. It follows from the general theory of the Legendre transform (see [1]) that F' and G' are inverse of each other and that $F(x)$ is the Legendre transform of $F^L(y)$. In other words, we have the classic formulas:

$$\left(F^L\right)^L = F$$

$$\left(F'\right)^{-1} = \left(F^L\right)'.$$

There is a well-defined theory of Legendre transform for convex functions. In that case, $\Omega = X$ and the Legendre formula is replaced by the Fenchel formula:

$$F^*(y) = \sup_x \{yx - F(x)\}$$

so that differentiability is no longer required. The Fenchel transform F^* will coincide with the Legendre transform F^L provided F is C^1 and strictly convex. We now proceed to give other classes of functions which have a well-defined Legendre transform. Roughly speaking, these will be functions $F(x)$, where x is a matrix and F depends only on the cofactors of x.

Given a number K and some $k \leq K$, we shall denote by $\mathscr{P}_K$ the set of strictly increasing maps of $\{1,...,k\}$ into $\{1,...,K\}$:

$$\mathscr{P}_K = \{\pi : \{1,...,k\} \to \{1,...,K\} \mid \pi(1) < ... < \pi(k)\}$$

and by $c(K,k) = C_K^k$ its cardinal. We can also think of $\mathscr{P}_K$ as being the set of ordered subsets of $\{1,...,K\}$ with k elements. Similar notations will hold for $\mathscr{P}_N$ and $c(k,N)$, provided $k \leq N$.

Consider a fixed N-dimensional space E. An element $x \in E^*$ will have coordinates x_n, $1 \leq n \leq N$. Given a family of K linear forms, $\left(x^1,...,x^K\right)$, and a number $k \leq \min\{K,N\}$, there are $c(K,k) \times c(N,k)$ square $k \times k$ matrices which can be extracted from the matrix x_n^k. Each of them is specified by a certain choice of k lines and k columns, that is by some $\pi \in \mathscr{P}_K$ and some $\sigma \in \mathscr{P}_K$. We shall denote it by:

$$x_\sigma^\pi = \left[x_{\sigma(j)}^{\pi(i)}\right]_{1\leq j\leq k}^{1\leq i\leq k}$$

and we shall denote by $\Delta_k^{\pi,\sigma}(x^1,...,x^K)$its determinant:

$$\Delta_k^{\pi,\sigma}(x^1,...,x^K) = \det[x_\sigma^\pi].$$

We then define a map $\Delta_k : (E^*)^K \to R^{c(K,k)} \times R^{c(N,k)}$ by:

$$\Delta_k = \left(\Delta_k^{\pi,\sigma}\right)_{\sigma\in\mathscr{P}_N}^{\pi\in\mathscr{P}_K}. \tag{2}$$

So the map Δ_k just associates with a $N \times K$ matrix $\left(x_n^k\right)$ the determinants of all the $k \times k$ matrices which can be extracted from it, that is, its k-cofactors.

Lemma 1 *We have:*

$$\Delta_k^{\pi,\sigma} = k \sum_{n,j} x_n^j \frac{\partial \Delta_k^{\pi,\sigma}}{\partial x_n^j}.$$

Proof This is just the Euler identity for k-homogeneous functions. □

Lemma 2 *Set:*

$$z(\pi,\sigma)_j^n = \frac{\partial}{\partial x_j^n} \det[x_\sigma^\pi].$$

We then have:

$$\det\left[z(\pi,\sigma)_\sigma^\pi\right] = \left(\det[x_\sigma^\pi]\right)^{k-1}.$$

Proof We know that $(z^{\pi}_{\sigma})^n_j$ is 0 if n does not belong to the image of π, or if j does not belong to the image of σ, and that otherwise it is just the cofactor of x^n_j in the matrix (x^{π}_{σ}). The last identity then follows from the well-known fact that the determinant of a $k\times k$ matrix raised to the $(k-1)$-th power is the determinant of the cofactor matrix. □

As we stated in the beginning, we are interested in functions of the $\left(x^k_n\right)$ which involve the Δ_k. We now make this idea precise. Consider the map:

$$\Delta = (\Delta_1,...,\Delta_K) : R^{NK} \to H$$

$$H = R^{NK} \times ... \times R^{c(N,k)c(K,k)} \times ... \times R^{c(N,K)}.$$

A function $\Phi : H \to R$ will be called *trivial* if it depends on the NK first coordinates only that is, if it factors through $(E^*)^K$. Note that every function $F : R^{NK} \to R$ can be written $F = \Phi\circ\Delta$, where $\Phi : H \to R$ is the identity on $(E^*)^K$ and sends all the other coordinates to 0. This is called the trivial factorisation.

Definition 2 A function $F : R^{NK} \to R$ will be called *k-adapted* if it factors through Δ_k, that is, if we have $F = \Phi\circ\Delta_k$ for some function $\Phi : R^{c(N,k)c(K,k)} \to R$. It is *adapted* if it factors non-trivially through Δ, that is, if we have $F = \Phi\circ\Delta$ for some non-trivial function $\Phi : H \to R$ such that $\Phi\circ\Delta_k$.

We now write the formula for the Legendre transform. Take an adapted function $F : R^{NK} \to R$:

$$F(x) = F\left(x^1,...,x^K\right) = F\left(x^j_n\right) = \Phi\left(\Delta^{\pi,\sigma}_k\right) = \Phi(\Delta)$$

and pair $K\times N$ matrices with $N\times K$ matrices by:

$$< x,y >= \sum_{n,j} x^j_n y^n_j .$$

Substitute in the definition (1):

$$\begin{aligned}
\Gamma_L(y_1,...,y_K) &= \left\{\sum_{n,j} x^j_n y^n_j - F\left(x^1,...,x^K\right) \mid y^n_j = \frac{\partial F}{\partial x^j_n}\right\} \\
&= \left\{\sum_{n,j} x^j_n y^n_j - \Phi(\Delta) \mid y^n_j = \sum_{k,\pi,\sigma} \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n}\right\} \\
&= \left\{\sum_{n,j,k,\pi,\sigma} x^j_n \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n} - \Phi(\Delta) \mid y^n_j = \sum_{k,\pi,\sigma} \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n}\right\} \qquad (3) \\
&= \left\{\sum_{k,\pi,\sigma} \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \sum_{n,j} x^j_n \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n} - \Phi(\Delta) \mid y^n_j = \sum_{k,\pi,\sigma} \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n}\right\} \\
&= \left\{\sum_{k,\pi,\sigma} k \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \Delta^{\pi,\sigma}_k - \Phi(\Delta) \mid y^n_j = \sum_{k,\pi,\sigma} \frac{\partial \Phi}{\partial \Delta^{\pi,\sigma}_k} \frac{\partial \Delta^{\pi,\sigma}_k}{\partial x^j_n}\right\}
\end{aligned}$$

where Δ stands for $\Delta\left(x^1,...,x^K\right)$. We rewrite the result in more compact notation:

$$\Gamma_L(y_1,...,y_K)=\left\{\sum_k k\frac{\partial\Phi}{\partial\Delta_k}\Delta_k-\Phi(\Delta)\mid y_j^n=\sum_k\frac{\partial\Phi}{\partial\Delta_k}\frac{\partial\Delta_k}{\partial x_n^j}\right\}. \tag{4}$$

We would like to give general conditions on Φ which would ensure that the right-hand side is a singleton, so that F has a well-defined Legendre transform. In addition, we would like to show that if F is k-adapted, then F^L is k-adapted as well. Unfortunately, we have not been able to fulfil this program (the calculations very quickly become horrendous) so we will be content with two examples.

3 Functions of the determinant

We take $N=K$. So let x be the square matrix with coefficients x_n^k. Denote by X_n^k the cofactor of x_n^k in X. We consider functions $F:R^N\to R$ of the following type:

$$F(x)=\Phi(\det x)$$

where $\Phi:R\to R$ is a C^1 function.

Let us apply the preceding theory. We have:

$$y_k^n=\frac{\partial F}{\partial x_n^k}=\Phi'(\det x)\frac{\partial\det x}{\partial x_n^k}=\Phi'(\det x)\det X_n^k. \tag{5}$$

Hence:

$$\sum_{n,k}y_k^nx_n^k=\Phi'(\det x)\sum_{n,k}x_n^k\det X_n^k=N\Phi'(\det X)\det X. \tag{6}$$

On the other hand, denoting by Y the matrix with coefficients y_k^n, and by z the matrix with coefficients $z_k^n=\det X_n^k$, we have:

$$\det Y=\left(\Phi'(\det x)\right)^N\det z=\Phi'(\det x)^N(\det x)^{N-1}. \tag{7}$$

Proposition 1 *Assume that Φ is such that the function $t\to t^{N-1}\Phi'(t)^N$ is invertible on (a,b), and let $\psi:(a',b')\to R$ be its inverse. Set $\Omega=\{x\mid a<\det x<b\}$ and $\Sigma=\{y\mid a'<\det y<b'\}$. Then the function $F:\Omega\to R$ given by:*

$$F(x)=\Phi(\det x)$$

has a Legendre transform $F^L:\Sigma\to R$ given by:

$$F^L(y)=N\,\Phi'(\psi(\det y))\,\psi(\det y)\;-\Phi(\psi(\det y)). \tag{8}$$

Proof Equation (7) gives $\det x=\psi(\det y)$. Writing (6) and (7) back into formula (3), we get:

$$\begin{aligned}\Gamma_L(y)&=\left\{N\,\Phi'(\det x)\det x-\Phi(\det x)\mid y_j^n=\Phi'(\det x)\det X_n^k\right\}\\&=\left\{N\,\Phi'(\psi(\det y))\,\psi(\det y)-\Phi(\psi(\det y))\right\}\end{aligned}$$

yielding a unique value. □

3.1 Example 1

Take $\Phi(t) = \frac{N}{p}|t|^{p/N}$, with $p \in R$, so that

$$F(x) = \frac{N}{p}|\det x|^{p/N}. \tag{9}$$

Note that F is homogeneous of degree p. If $p \neq 0$, then $t^{N-1}\Phi'(t)^N = t^{p-1}$, which is invertible provided $p \neq 1$, yielding $\psi(s) = s^{1/(p-1)}$. Substituting in the above, and taking advantage of the fact that Φ is homogeneous of degree p/N, we get:

$$F^L(y) = (p-1)\Phi(\psi(\det y)) = (p-1)\frac{N}{p}|\det y|^{\frac{p}{p-1}\frac{1}{N}}.$$

Proposition 2 *If $p \neq 0$ and $p \neq 1$, the function $F : R^{N^2} \to R$ defined by (9) has a Legendre transform $F^L : R^{N^2} \to R$ defined by:*

$$F^L(y) = \frac{N}{q}|\det y|^{q/N},$$

with $\frac{1}{p} + \frac{1}{q} = 1$.

Note that this duality holds for any value of p different from 0 and $1/N$, including negative ones. Note also that if $N > 1$ (the only interesting case), this duality between $\Phi(\det X)$ and $\Psi(\det Y)$ has nothing to do with convexity. On the one hand, if $p/N > 1$, so that $\Phi(t)$ is convex, then $0 < q/N < 1$, so that $\Psi(t)$ is not convex. On the other hand, it is easy to check that if $y \neq 0$, we can find a matrix $\bar{x}$ such that $\sum x_n^k y_k^n = 1$ and $\det \bar{x} = 0$; it follows that the function

$$x \to \sum x_n^k y_k^n - \frac{1}{p}(\det X)^p$$

is unbounded from above and from below (consider the sequences $x_n = \pm n\bar{x}$), and the critical point in the definition of the Legendre transform cannot be a global minimum or maximum.

3.2 Example 2

Take $p = 1$ in the above, so that $\Phi(t) = Nt^{1/N}$ and:

$$F(x) = N|\det x|^{1/N}. \tag{10}$$

The function F is defined on the whole of R^{N^2}. On the other hand, we have

$$y_k^n = \frac{\partial F}{\partial x_n^k}(x) = \Phi'(\det x)\det X_n^k$$

by (5), and $\det y = 1$ by (7). Consequently F' maps R^{N^2} onto the set $\Sigma = \{y \mid \det y = 1\}$. On the other hand, formula (8) yields quite simply $\Gamma_L(y) = 0$. Hence:

Proposition 3 *The Legendre transform of the function $F : R^{N^2} \to R$ given by (10) is the function $F^L : \Sigma \to R$ given by $F^L(y) \equiv 0$.*

3.3 Example 3

Take $\Phi(t) = \ln|t|$, so that:

$$F(x) = \ln|\det x|. \tag{11}$$

Set $\Omega = \{x \mid \det x \neq 0\}$.

Proposition 4 *The function $F : \Omega \to R$ defined by (11) has a Legendre transform $F^L : \Omega \to R$ defined by:*

$$F^L(y) = N + \ln|\det y|.$$

The proof is left to the reader. It follows that the function $G(x) = \ln|\det x| + N/2$ is self-dual, i.e. $G^L = G$.

4 The case of $(N-1)$-cofactors

We shall work with $N = K = 3$. We presume that similar results hold in the general case, but we have not been able to handle the notations.

Denote by $x$the 3×3 matrix with coefficients x_n^k, with $1 \le k \le K$ and $1 \le n \le N$. Denote by X_n^k the cofactor of x_n^k in x, and by Δ_n^k its determinant. Set

$$\Delta = \left(\Delta_n^k\right)_{1\le n\le N}^{1\le k\le K} \in R^9.$$

Let $\Phi : R^9 \to R$ be given. Consider the function:

$$F(x) = \Phi(\Delta). \tag{12}$$

The formula for the Legendre transform then becomes:

$$\Gamma_L(y) = \left\{ 2\sum_{n,k} \frac{\partial \Phi}{\partial \Delta_n^k}(\Delta)\Delta_n^k - \Phi(\Delta) \mid y_j^n = \sum_{n,k} \frac{\partial \Phi}{\partial \Delta_n^k}\frac{\partial \Delta_n^k}{\partial x_n^j} \right\}. \tag{13}$$

Let us simplify this formula a little bit by setting:

$$\Phi_k^n = \frac{\partial \Phi}{\partial \Delta_n^k}(\Delta).$$

We then have:

$$y_k^n = \sum_{i,j} \Phi_j^i \frac{\partial \Delta_i^j}{\partial x_n^k} = \sum_{i\neq n,\ j\neq k} \Phi_j^i \frac{\partial \Delta_i^j}{\partial x_n^k}(x) \tag{14}$$

because, if $k = j$ or $n = i$, the variable x_n^k does not appear in the cofactor X_n^k. If $k \neq j$, we shall denote by $p(j,k)$ the number in $\{1,2,3\}$ which is different from both k and j. Similarly, if $n \neq i$, we shall denote by $q(i,n)$ the number in $\{1,2,3\}$ which is different from both n and i. If $k \neq j$ and $n \neq i$, we have:

$$\Delta_i^j = (-1)^{m(i,j,k,n)} \left(x_n^k x_{q(i,n)}^{p(j,k)} - x_{q(i,n)}^k x_n^{p(j,k)} \right)$$

where: the exponent $m(i,j,k,n)$ is 0 if $k > p(j,k)$ and $n > q(i,n)$, or if $k < p(j,k)$ and $n < q(i,n)$, and $m(i,j,k,n) = 0$ otherwise. It follows that:

$$\frac{\partial \Delta_i^j}{\partial x_n^k} = (-1)^{m(i,j,k,n)} x_{q(i,n)}^{p(j,k)}$$

and hence:

$$y_k^n = \sum_{i\neq n\; j\neq k} (-1)^{(k-p(j,k))(n-q(i,n))} \Phi_j^i x_{q(i,n)}^{p(j,k)}.$$

Let us now consider the 3×3 matrix y with coefficients y_k^n, denote by Y_k^n the cofactor of y_k^n and by D_k^n its determinants. Without loss of generality, we can assume that $n = k = 1$, and we get:

$$\begin{aligned}
D_1^1 &= y_2^2 y_3^3 - y_3^2 y_2^3 \\
&= \left(\Phi_1^1 x_3^3 - \Phi_3^1 x_3^1 - \Phi_1^3 x_1^3 + \Phi_3^3 x_1^1\right)\left(\Phi_1^1 x_2^2 + \Phi_2^1 x_2^1 + \Phi_1^2 x_1^2 + \Phi_2^2 x_1^1\right) - \\
&\quad - \left(-\Phi_1^1 x_3^2 - \Phi_2^1 x_3^1 + \Phi_1^3 x_1^2 + \Phi_2^3 x_1^1\right)\left(-\Phi_1^1 x_2^3 + \Phi_3^1 x_2^1 - \Phi_1^2 x_1^3 + \Phi_3^2 x_1^1\right) \\
&= \left(\Phi_1^1\right)^2 \left(x_3^3 x_2^2 - x_3^2 x_2^3\right) + \left(\Phi_1^1 \Phi_2^1\right)\left(x_3^3 x_2^1 - x_3^1 x_2^3\right) + \\
&\quad + \left(\Phi_1^1 \Phi_1^2\right)\left(x_3^3 x_1^2 - x_3^2 x_1^3\right) + \left(\Phi_1^1 \Phi_2^2 x_3^3 x_1^1 - \Phi_2^1 \Phi_1^2 x_3^1 x_1^3\right) + \\
&\quad + \left(\Phi_3^1 \Phi_1^1\right)\left(-x_3^1 x_2^2 + x_3^2 x_2^1\right) + \left(-\Phi_3^1 \Phi_1^2 x_3^1 x_1^2 + \Phi_1^1 \Phi_3^2 x_3^2 x_1^1\right) + \\
&\quad + \left(-\Phi_3^1 \Phi_2^2 x_3^1 x_1^1 + \Phi_2^1 \Phi_3^2 x_3^1 x_1^1\right) + \left(\Phi_1^3 \Phi_1^1\right)\left(-x_1^3 x_2^2 + x_1^2 x_2^3\right) + \\
&\quad + \left(-\Phi_1^3 \Phi_2^1 x_1^3 x_2^1 + \Phi_2^3 x_1^1 \Phi_1^1 x_2^3\right) + \left(-\Phi_1^3 \Phi_2^2 x_1^3 x_1^1 + \Phi_2^3 x_1^1 \Phi_1^2 x_1^3\right) + \\
&\quad + \left(\Phi_3^3 x_1^1 \Phi_1^1 x_2^2 - \Phi_1^3 x_1^2 \Phi_3^1 x_2^1\right) + \left(\Phi_3^3 x_1^1 \Phi_2^1 x_2^1 - \Phi_2^3 x_1^1 \Phi_3^1 x_2^1\right) + \\
&\quad + \left(\Phi_3^3 x_1^1 \Phi_1^2 x_1^2 - \Phi_1^3 x_1^2 \Phi_3^2 x_1^1\right) + \left(\Phi_3^3 x_1^1 \Phi_2^2 x_1^1 - \Phi_2^3 x_1^1 \Phi_3^2 x_1^1\right).
\end{aligned}$$

Lemma 3 *If the matrix Φ_k^n has rank 1, then the D_k^n can be expressed in terms of the Δ_n^k as follows:*

$$\begin{aligned}
D_k^n = \Phi_k^n \big(&\Phi_3^2 \Delta_2^3 + \Phi_1^3 \Delta_3^1 + \Phi_2^3 \Delta_3^2 + \\
&+ \Phi_3^3 \Delta_3^3 + \Phi_3^2 \Delta_2^3 + \Phi_1^3 \Delta_3^1 + \Phi_2^3 \Delta_3^2 + \Phi_3^3 \Delta_3^3\big).
\end{aligned} \tag{15}$$

Proof If the matrix Φ_k^n has rank 1, all its 2-cofactors vanish, so that $\Phi_k^n \Phi_j^i = \Phi_j^n \Phi_k^i$. The previous expression then simplifies:

$$\begin{aligned}
D_1^1 = &\left(\Phi_1^1\right)^2 \Delta_1^1 + \left(\Phi_1^1 \Phi_2^1\right)\Delta_1^2 + \left(\Phi_1^1 \Phi_1^2\right)\Delta_2^1 + \\
&+ \left(\Phi_1^1 \Phi_2^2\right)\Delta_2^2 + \left(\Phi_3^1 \Phi_1^1\right)\Delta_1^3 + \left(\Phi_1^1 \Phi_3^2\right)\Delta_2^3 + \\
&+ \left(\Phi_1^3 \Phi_1^1\right)\Delta_3^1 + \left(\Phi_2^3 \Phi_1^1\right)\Delta_3^2 + \left(\Phi_3^3 \Phi_1^1\right)\Delta_3^3
\end{aligned}$$

and Φ_1^1 factors out. □

If Φ is homogeneous of degree α, the expression (15) simplifies by the Euler identity:

$$D_k^n = \alpha \Phi_k^n \Phi = \alpha \Phi(\Delta) \frac{\partial \Phi}{\partial \Delta_n^k}(\Delta) \tag{16}$$

and the formula (13)for the Legendre transform Γ_L of F becomes:

$$\Gamma_L(y) = \left\{(2\alpha - 1)\Phi(\Delta) \mid D = \alpha \Phi(\Delta)\Phi'(\Delta)\right\}.$$

Proposition 5 *Assume that the function $F : R^9 \to R$ is given by $F(x) = \Phi(\Delta)$, where Φ is homogeneous of degree α, and the matrix $\partial\Phi/\partial\Delta_n^k$ has rank 1 everywhere. Assume that $\Sigma = F'\left(R^9\right)$ is a submanifold, and that:*

$$\left[D_1 = \alpha\Phi(\Delta_1)\,\Phi'(\Delta_1) \text{ and } D_2 = \alpha\Phi(\Delta_2)\,\Phi'(\Delta_2)\right] \Longrightarrow \Phi(\Delta_1) = \Phi(\Delta_2).$$

Then F has a Legendre transform $F^L : \Sigma \to R$ given by

$$F^L(y) = \Psi(D) \tag{17}$$

where the D_k^n are the determinants of the 2-cofactors of y, and $\Psi(D) = (2\alpha - 1)\,\Phi(\Delta)$ for any D such that $D = \alpha\Phi(\Delta)\,\Phi'(\Delta)$.

4.1 Example 4

We consider functions $F : R^9 \to R$ of the following type:

$$F(x) = \left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^\beta = \Phi(\Delta).$$

We have

$$\Phi_k^i = \frac{\partial\Phi}{\partial\Delta_i^k} = \alpha\beta\left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^{\beta-1}(\sum_n \Delta_n^k)^{\alpha-1}$$

so clearly the matrix Φ_k^n has rank 1. The equations (16) become:

$$D_k^i = \alpha\beta\Phi_k^i\Phi = (\alpha\beta)^2\left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^{2\beta-1}(\sum_n \Delta_n^k)^{\alpha-1}$$

from which we get

$$D_k^1 = D_k^2 = D_k^3 \text{ for } k = 1,2,3. \tag{18}$$

In other words, the Legendre transform will live on the 3-dimensional subspace Σ of R^9 defined by the equations (18). Setting $D_k^i = D_k$ for every i, we continue the computations:

$$(D_k)^{\frac{\alpha}{\alpha-1}} = (\alpha\beta)^{\frac{2\alpha}{\alpha-1}}\left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^{(2\beta-1)\frac{\alpha}{\alpha-1}}(\sum_n \Delta_n^k)^\alpha,$$

$$\sum_k (D_k)^{\frac{\alpha}{\alpha-1}} = (\alpha\beta)^{\frac{2\alpha}{\alpha-1}}\left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^{\frac{2\alpha\beta-1}{\alpha-1}},$$

$$\left(\sum_k (D_k)^{\frac{\alpha}{\alpha-1}}\right)^{\frac{\beta(\alpha-1)}{2\alpha\beta-1}} = (\alpha\beta)^{\frac{2\alpha\beta}{2\alpha\beta-1}}\left((\sum_n \Delta_n^1)^\alpha + (\sum_n \Delta_n^2)^\alpha + (\sum_n \Delta_n^3)^\alpha\right)^\beta.$$

Finally, the Legendre transform of F turns out to be the function:

$$F^L(y) = (2\alpha\beta - 1)(\alpha\beta)^{-\frac{2\alpha\beta}{2\alpha\beta-1}} \left(\sum_k (D_k)^{\frac{\alpha}{\alpha-1}}\right)^{\frac{\beta(\alpha-1)}{2\alpha\beta-1}}$$

restricted to the 3-dimensional subspace $\Sigma \subset R^9$ defined by the relations $D_k^n = D_k$. Here, D_k^n denotes the cofactor of y_k^n in the matrix Y.

Note that F is homogeneous of degree $2\alpha\beta$ and F^L is homogeneous of degree $2\alpha\beta/(2\alpha\beta-1)$. Setting $p = 2\alpha\beta$ and $q = 2\alpha\beta/(2\alpha\beta-1)$, we find that;

$$\frac{1}{p} + \frac{1}{q} = 1.$$

4.2 Example 5

Let (x^1, x^2) be a pair of vectors in R^3. We consider functions $F : R^6 \to R$ of the following type:

$$F(x^1, x^2) = \Phi\left(\det\begin{vmatrix} x_1^1 & x_1^2 \\ x_2^1 & x_2^2 \end{vmatrix}, \det\begin{vmatrix} x_1^1 & x_1^2 \\ x_3^1 & x_3^2 \end{vmatrix}, \det\begin{vmatrix} x_2^1 & x_2^2 \\ x_3^1 & x_3^2 \end{vmatrix}\right)$$

where $\Psi : R^3 \to R$ is a C^1 function. In the previous framework, this can be understood as a function $F(X)$, where X is a 3×3 matrix, which depends only on the first three cofactors. Clearly the rank condition will hold, and the previous results apply. It will be more convenient, however, to run through the computations again in that particular case, with simplified notations.

Set $\Delta = (\Delta_3, \Delta_2, \Delta_1)$, with:

$$\Delta_3 = x_1^1 x_2^2 - x_1^2 x_2^1,$$
$$\Delta_2 = x_1^1 x_3^2 - x_1^2 x_3^1,$$
$$\Delta_1 = x_2^1 x_3^2 - x_2^2 x_3^1.$$

We have $F(x^1, x^2) = \Phi(\Delta)$. Set $y_k^n = \partial F / \partial x_n^k$ and compute the cofactors. We get:

$$D^3 = y_1^1 y_2^2 - y_2^1 y_1^2 = \left(\frac{\partial\Phi}{\partial\Delta_3}\right)^2 \Delta_3 + \left(\frac{\partial\Phi}{\partial\Delta_3}\frac{\partial\Phi}{\partial\Delta_1}\right)\Delta_1 + \left(\frac{\partial\Phi}{\partial\Delta_2}\frac{\partial\Phi}{\partial\Delta_3}\right)\Delta_2,$$

$$D^2 = y_1^1 y_2^3 - y_2^1 y_1^3 = \left(\frac{\partial\Phi}{\partial\Delta_2}\right)^2 \Delta_2 + \left(\frac{\partial\Phi}{\partial\Delta_3}\frac{\partial\Phi}{\partial\Delta_2}\right)\Delta_3 + \left(\frac{\partial\Phi}{\partial\Delta_2}\frac{\partial\Phi}{\partial\Delta_1}\right)\Delta_1,$$

$$D^1 = y_1^2 y_2^3 - y_2^2 y_1^3 = \left(\frac{\partial\Phi}{\partial\Delta_1}\right)^2 \Delta_1 + \left(\frac{\partial\Phi}{\partial\Delta_1}\frac{\partial\Phi}{\partial\Delta_2}\right)\Delta_2 + \left(\frac{\partial\Phi}{\partial\Delta_1}\frac{\partial\Phi}{\partial\Delta_3}\right)\Delta_3.$$

We summarize:

$$D^n = \frac{\partial\Phi}{\partial\Delta_n}\left[\frac{\partial\Phi}{\partial\Delta_1}\Delta_1 + \frac{\partial\Phi}{\partial\Delta_2}\Delta_2 + \frac{\partial\Phi}{\partial\Delta_3}\Delta_3\right],\ n = 1,2,3. \tag{19}$$

As a particular case, consider the function:

$$F\left(x^1,x^2\right)=\left[\left(x_1^1x_2^2-x_1^2x_2^1\right)^{\alpha}+\left(x_1^1x_3^2-x_1^2x_3^1\right)^{\alpha}+\left(x_2^1x_3^2-x_2^2x_3^1\right)^{\alpha}\right]^{1/\beta}. \tag{20}$$

When $\alpha=1/\beta=2$, this gives the area of the triangle spanned by the vectors x^1 and x^2. We apply the preceding result, with:

$$\Phi\left(\Delta_3,\Delta_2,\Delta_1\right)=\left[(\Delta_1)^{\alpha}+(\Delta_2)^{\alpha}+(\Delta_3)^{\alpha}\right]^{\beta}.$$

The system (19) becomes:

$$D^n=(\alpha\beta)^2\left[(\Delta_1)^{\alpha}+(\Delta_2)^{\alpha}+(\Delta_3)^{\alpha}\right]^{2\beta-1}(\Delta_n)^{\alpha-1}$$

and can easily be inverted (note that if $\alpha=1/\beta=2$, we get the identity). We get:

$$\Delta_n=(\alpha\beta)^{-\frac{2\alpha-1}{2\alpha\beta-1}\frac{1}{\alpha-1}}\left[\left(D^1\right)^{\frac{\alpha}{\alpha-1}}+\left(D^2\right)^{\frac{\alpha}{\alpha-1}}+\left(D^3\right)^{\frac{\alpha}{\alpha-1}}\right]^{-\frac{2\beta-1}{2\alpha\beta-1}}(D^n)^{\frac{1}{\alpha-1}}.$$

Substituting into formula (17), and taking advantage of the fact that Φ is homogeneous of degree $\alpha\beta$, we get the Legendre transform:

$$\begin{aligned}
F^L(y_1,y_2)&=(2\alpha\beta-1)\left[(\Delta_1)^{\alpha}+(\Delta_2)^{\alpha}+(\Delta_3)^{\alpha}\right]^{\beta}\\
&=(2\alpha\beta-1)(\alpha\beta)^{-\frac{2\alpha-1}{2\alpha\beta-1}\frac{\alpha\beta}{\alpha-1}}\left[\left(D^1\right)^{\frac{\alpha}{\alpha-1}}+\left(D^2\right)^{\frac{\alpha}{\alpha-1}}+\left(D^3\right)^{\frac{\alpha}{\alpha-1}}\right]^{\frac{\alpha-1}{2\alpha\beta-1}\beta}\\
&=(2\alpha\beta-1)(\alpha\beta)^{-\frac{2\alpha-1}{2\alpha\beta-1}\frac{\alpha\beta}{\alpha-1}}\times\\
&\quad\times\left[\left(y_1^2y_2^3-y_2^2y_1^3\right)^{\frac{\alpha}{\alpha-1}}+\left(y_1^1y_2^3-y_2^1y_1^3\right)^{\frac{\alpha}{\alpha-1}}+\left(y_1^1y_2^2-y_2^1y_1^2\right)^{\frac{\alpha}{\alpha-1}}\right]^{\frac{\alpha-1}{2\alpha\beta-1}\beta}.
\end{aligned}$$

Note that if $\alpha=1/\beta=2$, we find $F=F^L$: the function F is its own Legendre transform. Note also that F is homogeneous of degree $2\alpha\beta$ and F^L homogeneous of degree $2\alpha\beta/(2\alpha\beta-1)$. Setting $p=2\alpha\beta$ and $q=2\alpha\beta/(2\alpha\beta-1)$, we find that

$$\frac{1}{p}+\frac{1}{q}=1$$

as before.

5 A variational problem

As a example of possible application of this kind of duality, let us consider the following problem. Given a positive definite quadratic form (Ax,x) on R^{N^2}, and a $N\times N$-matrix f, we want to solve $\Phi'(x)=0$, where::

$$\Phi(x)=\frac{1}{2}(Ax,x)-\frac{N}{p}|\det x|^{p/N}-(f,x).$$

Such points are called critical points of Φ. Any critical point of F solves the system:

$$Ax=|\det x|^{p/N-1}X+f$$

where X is the matrix of cofactors of x.

Proposition 6 *If $p<2$, there is at least one critical point for Φ.*

Proof Since $p<2$, the function Φ is coercive: $\Phi(x)\to\infty$ when $\|x\|\to\infty$. So it attains its minimum at some x, which has to be a critical point. □

We now use duality theory to treat the case $p>2$. We shall use an extension of the Clarke duality formula (see [2]):

Proposition 7 *Suppose $F(x)$ has a Legendre transform $F^L(y)$. Consider the functions Φ and Ψ defined by:*

$$\Phi(x)=\frac{1}{2}(Ax-f,x)-F(x)$$

$$\Psi(y)=\frac{1}{2}(Ay+f,y)-F^L(Ay).$$

If y is a critical point of Ψ, then $x=y+A^{-1}f$ is a critical point of Φ.

Proof If y is a critical point of Ψ, we have $\Psi'(y)=0$, and hence $Ay+f=A\left(F^L\right)'(Ay)$. Since A is invertible, it follows that $y+A^{-1}f=\left(F^L\right)'(Ay)$. Since $\left[\left(F^L\right)'\right]^{-1}=F'$, it follows that

$$F'\left(y+A^{-1}f\right)=Ay=A\left(y+A^{-1}f\right)-f$$

so that $F'(x)=Ax-f$, as desired. □

Proposition 8 *If $p>2$, there is at least one non-trivial critical point for Φ.*

Proof Consider the function:

$$\Psi(y)=\frac{1}{2}(Ay+f,y)-\frac{N}{q}|\det Ay|^{q/N}.$$

By proposition 6, the function Ψ has a critical point $\bar{y}\neq 0$. By proposition 7 it is also a critical point of Φ. □

Note that in this case $\inf\Phi=-\infty$, so that the critical point x cannot be a minimizer.

Acknowledgements The author thanks Jey Sivaloganathan for discussing an early version of the paper and Paolo Marcellini for suggesting Example 3.

References

1. Ekeland, I.: Duality in non-convex optimization and calculus of variations. SIAM J. Opt. and Control **15**, 905-934 (1977)
2. Ekeland, I.: Convexity methods in Hamiltonian mechanics. Springer-Verlag, Berlin (1990)