Working Paper Series
Department of Economics
University of Verona

A mapping associated to a quadratic optimization problem with linear constraints

Alberto Peretti

WP Number: 22 December 2014

ISSN: 2036-2919 (paper), 2036-4679 (online)

A mapping associated to a quadratic optimization problem with linear constraints

Alberto Peretti
Department of Economics, University of Verona
Verona, Italy
e-mail: alberto.peretti@univr.it

December, 2014

Abstract

In this working paper we go on with the study of a mapping in $\mathbb{R}^n$ associated to a quadratic optimization problem with one equality linear constraint. After showing some general properties related to homogeneity and the inverse mapping, we present some results regarding how the mapping behaves with norms of vectors. Then some aspects related either to invariant points or invariant subspaces are investigated.

Keywords. Constrained optimization, quadratic programming, linear mappings.

AMS Classification. 90C20, 90C46

JEL Classification. C61, C65

1 Introduction and motivations

In a previous working paper [1] a mapping related to a quadratic optimization problem was investigated. Consider the quadratic optimization problem with one linear constraint

$$\begin{array}{ll} \min_x & x^T A x \\ \text{s.t.} & a^T x = a_0 \end{array} \tag{1}$$

where $x, a \in \mathbb{R}^n$, $a_0 \in \mathbb{R}$ and A is a square $n \times n$ (symmetric) positive definite matrix. The necessary and sufficient optimality conditions[1] on the Lagrangian function

$$\mathcal{L}(x, \lambda) = x^T A x - \lambda(a^T x - a_0) \qquad (\lambda \in \mathbb{R}),$$

are

$$\begin{cases} 2Ax - \lambda a = 0 \\ a^T x = a_0. \end{cases}$$

We get

$$\begin{cases} x = \frac{\lambda}{2} A^{-1} a \\ \frac{\lambda}{2} a^T A^{-1} a = a_0, \end{cases}$$

and then

$$\lambda = \frac{2a_0}{a^T A^{-1} a} \qquad \text{and} \qquad x = a_0 \frac{A^{-1} a}{a^T A^{-1} a}. \tag{2}$$

Equation (2) gives the optimal solution for problem (1). The solution vector is proportional to vector $\frac{A^{-1}a}{a^T A^{-1} a}$ or equivalently to vector $A^{-1}a$. In the former case we may observe that the proportionality constant is given by the "constraint level" a_0. In the relation between the solution vector and the "initial data" of the minimization problem, in particular the constraint vector a, the role is performed by the mapping

$$x \mapsto \frac{A^{-1} x}{x^T A^{-1} x}.$$

In [wp2009] some properties of this mapping were investigated. We carry on presenting some other properties that have both an algebraic and geometric meaning.

1.1 Some properties

Consider the mapping $\varphi : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}$ defined by

$$\varphi(x) = \frac{Ax}{x^T A x}, \tag{3}$$

[1] Conditions are also sufficient for optimality as the problem is convex in the assumptions.

where A is a (symmetric) positive definite $n \times n$ matrix. The assumption on the matrix A is motivated by the optimization problem which the mapping takes its origin from, but some of the following results hold in the more general assumption that A is nonsingular.

Lemma 1 *If A is nonsingular then φ is a homogeneous function of degree -1, that is, if $\alpha \neq 0$ is a scalar then*

$$\varphi(\alpha x) = \frac{1}{\alpha}\varphi(x).$$

Proof. We have

$$\text{for every } \alpha \neq 0 \quad \varphi(\alpha x) = \frac{\alpha A x}{\alpha^2 x^T A x} = \frac{1}{\alpha}\varphi(x). \qquad \square$$

Proposition 1 *If A is symmetric and nonsingular then φ is a one to one correspondence in $\mathbb{R}^n \setminus \{0\}$.*

Proof. We prove that the mapping

$$\psi(y) = \frac{A^{-1}y}{y^T A^{-1} y}$$

is the inverse of φ, that is $\varphi(\psi)$ is the identity function. We have

$$\begin{aligned} \varphi(\psi(y)) &= y^T A^{-1} y \; \varphi(A^{-1}y) \\ &= y^T A^{-1} y \; \frac{AA^{-1}y}{(A^{-1}y)^T AA^{-1}y} \\ &= y^T A^{-1} y \; \frac{y}{y^T A^{-1} y} = y. \end{aligned}$$

For the first equality the homogeneity has been used. Symmetry of A^{-1} is used at the end when $(A^{-1})^T = A^{-1}$. In the same way $\psi(\varphi(x)) = x$ can be proved. $\square$

Some geometric properties about how the mapping behaves with respect to vector norms are described.

In Figure 1 three numerical examples are shown. In particular for each case the set of norm 1 points is shown in black, its image through the linear mapping Ax is shown in blue and its image through the mapping φ is shown in red.

A diagonal matrix has been used in the three cases: $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ on the left, $A = \begin{pmatrix} 4/10 & 0 \\ 0 & 1 \end{pmatrix}$ in the middle and $A = \begin{pmatrix} 7/10 & 0 \\ 0 & 2/10 \end{pmatrix}$ on the right.

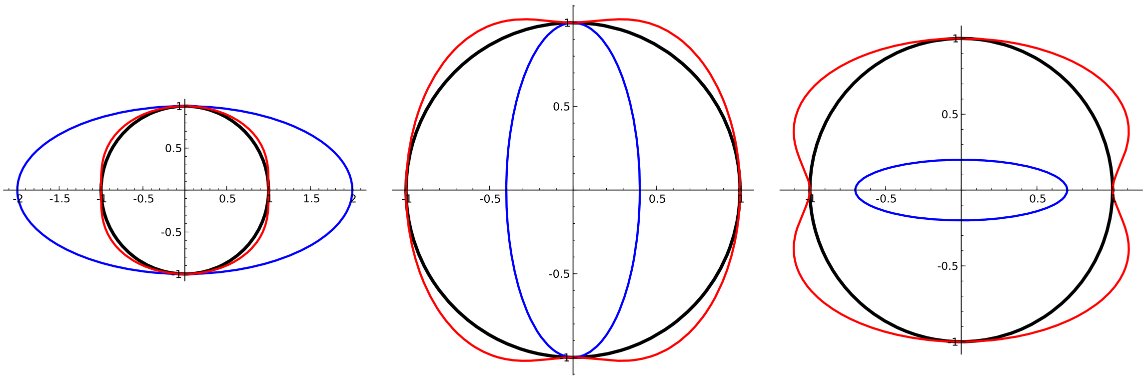

Figure 1: Examples in $\mathbb{R}^2$

Proposition 2 *If A is nonsingular then the inequality*

$$\|x\|\|\varphi(x)\| \geq 1 \tag{4}$$

holds for each $x \in \mathbb{R}^n \setminus \{0\}$.

Proof. From the Cauchy–Schwartz inequality

$$\|x\|\|\varphi(x)\| \geq |x^T\varphi(x)| = x^T\frac{Ax}{x^TAx} = 1. \qquad \square$$

Remark. It follows that if either $\|x\|$ or $\|\varphi(x)\|$ is less than or equal to 1, the other is greater than or equal to 1. In particular if $\|x\| \leq 1$ then $\|\varphi(x)\| \geq 1$. Also, in particular, if $\|x\| = 1$ then $\|\varphi(x)\| \geq 1$.

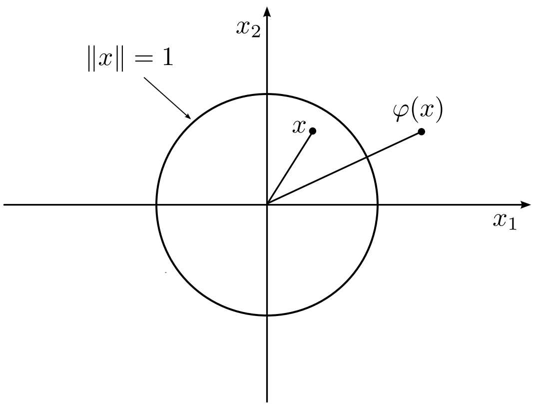

Figure 2: $\|x\| \leq 1 \ \Rightarrow \|\varphi(x)\| \geq 1$

Remark. If $\|x\|$ is equal to 1, $\|\varphi(x)\|$ can be greater than 1. Take for example $x = u^1$ (the first fundamental vector in $\mathbb{R}^n$); then

$$\|\varphi(x)\| = \frac{\|Au^1\|}{(u^1)^TAu^1} = \frac{\|a^1\|}{a_{11}}$$

where a^1 is the first column of the matrix A. This is greater than 1 when a^1 has at least one non zero element apart from a_{11}.

Remark. The implication $\|x\| \leq 1 \;\Rightarrow\; \|\varphi(x)\| \geq 1$ is not an "if and only if" condition, as the analogous implication that $\|x\| \geq 1$ takes to $\|\varphi(x)\| \leq 1$ cannot be derived from Proposition 2 and in fact is not true, as the previous remark already shows. But a stronger case with both $\|x\| > 1$ and $\|\varphi(x)\| > 1$ can be found.[2]

Remark. It may seem that inequality (4) together with the existence of the inverse, that satisfies the same inequality, could guarantee the if and only if condition. Though it is not true, as a simple example with an invertible real function shows. Consider the real function $f(x) = 1 + \frac{1}{x}$ in $(0, +\infty)$. For this function the inequality

$$xf(x) \geq 1 \quad \text{(equivalent to } x + 1 \geq 1\text{)}$$

holds. The inverse is $f^{-1}(y) = \frac{1}{y-1}$ and it satisfies the corresponding equality

$$yf^{-1}(y) \geq 1 \quad \text{(equivalent to } y \geq y - 1\text{)}.$$

But the implication $x \geq 1 \Rightarrow f(x) \leq 1$ is not true as $f(x)$ is greater than or equal to 1 for any value of x.

Remark. A similar if and only if condition holds instead if we introduce in the codomain a different "reference value" for the norm. Given any vector x, suppose $x = \alpha u$, where u is a unit vector. Let $\eta(x)$ be the norm of the image of the unit vector u through the mapping φ, so let's define

$$\eta(x) = \|\varphi(u)\|.$$

Then the following condition holds:

$$\|x\| \leq 1 \;\Leftrightarrow\; \|\varphi(x)\| \geq \eta(x).$$

Proof. Just consider that, if $x = \alpha u$, then $\|x\| \leq 1$ if and only if $\alpha \leq 1$. Moreover

$$\|\varphi(x)\| = \left\| \frac{1}{\alpha}\varphi(u) \right\| = \frac{1}{\alpha}\eta(u).$$

[2]Consider for example the matrix $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and let $x = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We have $\varphi(x) = \frac{1}{\alpha}\varphi\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3\alpha}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\|x\| = \alpha\sqrt{2}$ and $\|\varphi(x)\| = \frac{\sqrt{5}}{3\alpha}$. The system $\begin{cases} \alpha\sqrt{2} > 1 \\ \frac{\sqrt{5}}{3\alpha} < 1 \end{cases}$ is equivalent to $\begin{cases} \alpha > \frac{1}{\sqrt{2}} \\ \alpha < \frac{\sqrt{5}}{3} \end{cases}$ and it does have solutions.

Then the thesis follows. $\square$

The following Figure 3 shows the relationship between the norms of x and $\varphi(x)$. The red line represents the points that are images of unit vectors. Their norms give the values of the function η.

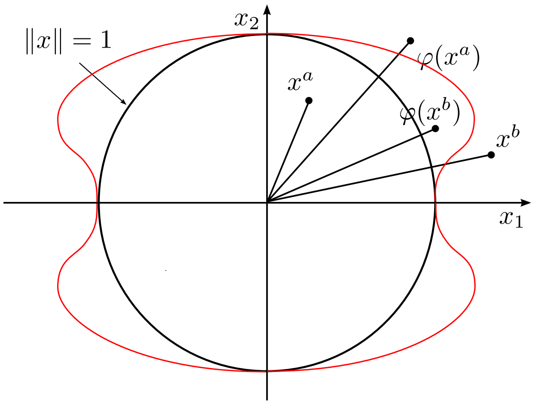

Figure 3: $\|x\| \leq 1 \;\Leftrightarrow \|\varphi(x)\| \geq \eta(x)$

Corollary 1 *If* $\|x\| \leq 1$ *then* $x^T Ax \leq \|Ax\|$.

Proof. If $\|x\| \leq 1$ from Proposition 2 $\|\varphi(x)\| \geq 1$, that is

$$\frac{\|Ax\|}{x^T Ax} \geq 1. \quad \text{Then } x^T Ax \leq \|Ax\|.$$

Let's consider now some particular cases.

Suppose A is the identity matrix. Then

$$\varphi(x) = \frac{x}{x^T x} = \frac{x}{\|x\|^2}.$$

Remark. It is a kind of non homogeneous normalization, something like a normalization with a negative degree of homogeneity.

Remark. If $\|x\| = 1$ then $\varphi(x) = x$, hence φ is the identity on the unit vectors. In this case $\|x\| \leq 1$ if and only if $\|\varphi(x)\| \geq 1$ as $\|\varphi(x)\| = \frac{1}{\|x\|}$ and then $\|x\|\|\varphi(x)\| = 1$.

Suppose A is a diagonal matrix, let's say $A = D$.

Remark. The fundamental vectors with their opposites are fixed points of the mapping φ. In fact

$$\varphi(u^i) = \frac{Du^i}{(u^i)^T Du^i} = \frac{d_i u^i}{d_i} = u^i$$

and for the opposites homogeneity gives the analogous result.

Remark. In general the unit vectors are not fixed points of the mapping. In fact take for example $x = \frac{1}{\sqrt{2}}(u^1 + u^2)$. We have

$$\begin{aligned} x^T Dx &= \frac{1}{2}(u^1 + u^2)^T D(u^1 + u^2) \\ &= \frac{1}{2}\left((u^1)^T Du^1 + (u^1)^T Du^2 + (u^2)^T Du^1 + (u^2)^T Du^2\right) \\ &= \frac{1}{2}(d_1 + d_2) \end{aligned}$$

and

$$\varphi(x) = \frac{Dx}{x^T Dx} = \frac{2}{d_1 + d_2} \cdot \frac{1}{\sqrt{2}} D(u^1 + u^2) = \frac{\sqrt{2}}{d_1 + d_2}(d_1 u^1 + d_2 u^2).$$

Then

$$\|\varphi(x)\| = \frac{\sqrt{2}}{d_1 + d_2}\|d_1 u^1 + d_2 u^2\| = \frac{\sqrt{2}}{d_1 + d_2}(d_1^2 + d_2^2)^{1/2}$$

and this is equal to $\frac{\sqrt{2}}{3} \cdot \sqrt{5} = \frac{\sqrt{10}}{3}$ if $d_1 = 2$ and $d_2 = 1$.

Remark. The particular case with a diagonal matrix is significant because if the matrix A is non singular and symmetric, by means of a change of variable and as a consequence of the spectral theorem we can always refer to the diagonal case, as we shall see shortly.

Let's consider now some properties related to the matrix eigenvalues and eigenvectors.

2 Role of the eigenvectors

Proposition 3 *If A is positive definite and u is a unit eigenvector of A then u is a fixed point of the mapping φ.*

Proof. Suppose λ is an eigenvalue of A and let's consider a unit eigenvector u in the associated eigenspace. Hence $Au = \lambda u$ and $\|u\| = 1$.
Then

$$\varphi(u) = \frac{Au}{u^T Au} = \frac{\lambda u}{u^T(\lambda u)} = \frac{u}{\|u\|^2} = u. \qquad \square$$

Remark. No need to say $\lambda \neq 0$ because the matrix A is positive definite and then has positive eigenvalues.

Remark. The opposites of unit eigenvectors are also fixed points for the mapping φ. This follows from homogeneity of the mapping. In the eigenspace associated to the unit eigenvector u u itself and its opposite are the only fixed points for the mapping. In fact from the homogeneity again

$$\varphi(\alpha u) = \frac{1}{\alpha}\varphi(u) = \frac{1}{\alpha}u.$$

Then the eigenspace is invariant through the mapping, but its single elements are not.

Now we want to prove that the only fixed points of the mapping φ are some of the eigenvectors of the matrix A. Let's state first that we can always change the coordinates system and get a diagonal case.

Let's set

$$y = \frac{Ax}{x^T Ax} \tag{5}$$

where x and y are the expressions of vectors in the canonical basis of $\mathbb{R}^n$.

Being A symmetric, from the spectral theorem, we know that an orthogonal matrix P exists such that $D = P^T AP$ is a diagonal matrix carrying as diagonal elements the eigenvalues of A. Consider the change of coordinates system given by the orthogonal matrix P and let's call $x_{\mathcal{P}}$ and $y_{\mathcal{P}}$ the expressions of x and y in the new basis $\mathcal{P}$. Then $x = Px_{\mathcal{P}}$ and $x_{\mathcal{P}} = P^T x$ are the equations of the change of variable and this is equivalent to suppose that the columns of P are the expressions in the canonical basis of the fundamental vectors in the basis $\mathcal{P}$. Moreover these columns are the matrix A eigenvectors.

If in (5) we change the coordinates system into the new basis $\mathcal{P}$ we get

$$Py_{\mathcal{P}} = \frac{APx_{\mathcal{P}}}{(Px_{\mathcal{P}})^T A(Px_{\mathcal{P}})} \qquad \text{that is} \qquad y_{\mathcal{P}} = \frac{P^T APx_{\mathcal{P}}}{x_{\mathcal{P}}^T P^T APx_{\mathcal{P}}}.$$

Then

$$y_{\mathcal{P}} = \frac{Dx_{\mathcal{P}}}{x_{\mathcal{P}}^T Dx_{\mathcal{P}}}.$$

The particular case with a diagonal matrix is nearly the only significant case as we can always reduce to this case with a convenient change of the coordinates system.

We have already showed that in the diagonal case the fundamental vectors together with their opposites are fixed points for the mapping φ.

Lemma 2 *Suppose $A = D$ is a diagonal matrix with distinct eigenvalues (diagonal elements) d_i. If v is a fixed point for the mapping $x \mapsto \frac{Dx}{x^T Dx}$, v must be either a fundamental vector or its opposite.*

Proof. Let's suppose v is a fixed point for the mapping. It means that

$$\frac{Dv}{v^T Dv} = v.$$

The vector equality above is equivalent to the system of equations

$$\frac{d_i v_i}{\sum_{j=1}^n d_j x_j^2} = v_i \qquad i = 1, \dots, n$$

that is

$$\left(\frac{d_i}{\sum_{j=1}^n d_j v_j^2} - 1 \right) v_i = 0 \qquad i = 1, \dots, n.$$

The terms in parenthesis on the left are distinct from the hypothesis that the d_i's are distinct. If they are all non zero, the solution is the null vector and it is not acceptable. Otherwise just one of them is zero, and then all the components of v are zero except for one, say v_i. The system is reduced to the single equation

$$\frac{d_i}{d_i v_i^2} = 1 \quad \text{that is} \quad v_i^2 = 1.$$

This means the only solutions are the i-th fundamental vector and its opposite. □

Lemma 3 *Suppose $A = D$ is a diagonal matrix with multiple eigenvalues. If v is a fixed point for the mapping $x \mapsto \frac{Dx}{x^T Dx}$, v must be a unit vector in the subspace spanned by some fundamental vectors.*

Proof. Let's suppose v is a fixed point for the mapping. It means that

$$\frac{Dv}{v^T Dv} = v$$

and again this is equivalent to the system of equations

$$\frac{d_i v_i}{\sum_{j=1}^n d_j x_j^2} = v_i \quad \text{that is} \quad \left(\frac{d_i}{\sum_{j=1}^n d_j v_j^2} - 1 \right) v_i = 0 \qquad i = 1, \dots, n.$$

Suppose there are multiple eigenvalues. Let's say we have the eigenvalues $d_{\nu_1}, d_{\nu_2}, \dots, d_{\nu_s}$, each one with its multiplicity $\mu_1, \mu_2, \dots, \mu_s$, one at least greater than 1. Moreover let's call I_k the set of indices i such that v_i is in the equation corresponding to the d_{ν_k} eigenvalue. Let's consider as before the terms in parenthesis

$$\frac{d_{\nu_k}}{\sum_{j=1}^n d_j v_j^2} - 1 \qquad k = 1, \dots, s.$$

If they are all non zero, the solution is the null vector and it is not acceptable. Otherwise just one of them is zero, suppose the one with $k = 1$, without loosing generality. Then all the v_i in the equations with $k \neq 1$ must be zero and the system is reduced to

$$\frac{d_{\nu_1}}{\sum_{j=1}^n d_j v_j^2} - 1 = 0 \quad \text{with} \quad v_i = 0 \text{ if } i \in \bigcup_{k \neq 1} I_k.$$

This is equivalent to

$$\frac{d_{\nu_1}}{\sum_{j \in I_1} d_{\nu_1} v_j^2} = 1 \quad \text{that is} \quad \sum_{j \in I_1} v_j^2 = 1.$$

These are the unit vectors in the subspace spanned by the fundamental vectors u^j with $j \in I_1$, that is the eigenspace associated to the eigenvalue d_{ν_1}.

Proposition 4 *The only fixed points of the mapping φ are the unit eigenvectors of the matrix A.*

Proof. It follows from the two previous lemmas. □

Remark. The aim for a further analysis of the mapping properties is first to go deeper into the relation between the initial data of a quadratic optimization problem and its optimal solution. Secondly, since this kind of optimization problem is the underlying model for the classical mean–variance approach to portfolio selection, the properties of the mapping could have an interesting meaning in the classical theory of efficient portfolio choices.

References

[1] A. Peretti, "On a particular mapping in $\mathbb{R}^n$", working paper n.68 of the Department of Economics, University of Verona, 2009.

[2] G.H. Golub, C.F. Van Loan, "Matrix computations", The Johns Hopkins University Press, Baltimore, 1987.

[3] J.M. Ortega, "Matrix theory", Plenum, New York, 1987

[4] A. Peretti, "On the selection of an efficient portfolio in the mean–variance approach", *Journal of Interdisciplinary Mathematics*, Vol. 7, No. 1, 2004.