ON CONVEX DECOMPOSITIONS

by J. Borwein

Dalhousie University

Abstract: Functions which are the differences of two convex functions are examined both in R^n and in locally convex spaces.

One of the central results on convex functions concerns the maximal monotonicity of their subgradients [4]. A natural question to ask is: what operators are the differences of maximal monotone operators? On the line, of course, these are exactly the functions of bounded variation. In general the question seems very hard, but it does suggest the related question of what functions are convex differences. An answer to this will of course provide information concerning the previous question and, in addition, the space thus generated will inherit many of the pleasant analytic properties of convex functions such as continuity on the relative interior of the domain of definition and differentiability densely, [4].

We make the following preliminary definitions. Let X be a locally convex topological vector space and let $\Omega \subset X$ be a compact convex subset of X. Let $C(\Omega)$ denote the continuous real valued functions on Ω with the sup norm and let $C_0(\Omega)$ denote the convex functions on Ω. We will denote $(C_0(\Omega) - C_0(\Omega)) \cap C(\Omega)$ by $K(\Omega)$.

<u>Proposition 1</u>: $K(\Omega)$ is a Riesz subspace of $C(\Omega)$. That is $K(\Omega)$ is a sublattice and subspace of $C(\Omega)$.

<u>Proof:</u> Let $e \in K(\Omega)$. Then there exist $a,b \in C_0(\Omega)$ with $e = a - b$. Then

$$e^- = \min(e,0) = a - \max(a,b)$$

and

$$e^+ = \max(e,0) = \max(a,b) - b .$$

Now max (a,b) is convex since it is the pointwise maximum of convex functions. It follows that e^+ and $e^- \in C_0(\Omega) - C_0(\Omega)$.

It is easy to verify that this is sufficient for $C_0(\Omega) - C_0(\Omega)$ to be a lattice. This in turn implies that $K(\Omega)$ is a sublattice of $C(\Omega)$. Since $C_0(\Omega)$ is a cone, $K(\Omega)$ is clearly a subspace. □

<u>Theorem 1:</u> $K(\Omega)$ is dense in $C(\Omega)$ with the supnorm.

<u>Proof:</u> It suffices by the Kakutani-Stone Theorem ([3]) to show that for any $f \in C(\Omega)$ and any x_1, x_2 in Ω and $\epsilon > 0$ there is some $g \in K(\Omega)$ with $|g(x_i) - f(x_i)| < \epsilon \quad i = 1,2$.

<u>Case (i):</u> x_1, x_2 are linearly independent. There then exist continuous linear functionals g_i on X with $g_i(x_j) = \delta_{ij} f(x_i)$. Then

$$g = g_1 - q_2 \in K(\Omega)$$

and $$g(x_i) = f(x_i) \quad \text{for} \quad i = 1,2 \ .$$

<u>Case (ii):</u> x_2 and x_1 are dependent linearly. We may assume that $x_2 = \lambda x_1$, $\lambda \in R$, $\lambda \neq 1$.

Let h be defined, on the subspace generated by x_1, by

$$h(tx_1) = (f(x_2) - f(x_1)) \left(\frac{t-\lambda}{\lambda-1}\right) + f(x_2) \ .$$

Then $h(x_1) = f(x_1)$ and $h(x_2) = f(x_2)$. Now h is continuous and affine from the subspace generated by x_1 and thus has a continuous extension to X, by the Hahn-Banach Theorem [1]. Thus in either case there is some convex (affine) function in $K(\Omega)$ with the same values as f . It follows that $\overline{K(\Omega)} = C(\Omega)$. □

Remark: It suffices that Ω be completely regular for Theorem 1 to hold with the topology of uniform convergence on compact sets. Note also that the Theorem promises that any function in $C(\Omega)$ can be approximated by functions which are differences of convex, continuous maps which is slightly stronger than by functions in $K(\Omega)$.

In R one can completely characterize $K(\Omega)$, for $\Omega = [a,b]$ (with $\Omega \subset \operatorname{int}(\operatorname{dom} f)$).

Theorem 2: $f \in K([a,b])$ if and only if f is absolutely continuous and $f = \int g$ with $g \in BV[a,b]$.

Proof: Since $f \in K(\Omega)$, f is the difference of convex functions f_i which, being convex, are Lipshitz on the relative interior of their domains [4]. Thus f is absolutely continuous. Moreover, since f_i is convex on $[a,b]$, $f_i'(x)$ exists except countably and is monotone. Let A_i be the domain of definition and let

$$g_i(x) = \sup_{y \in A_i, y \leq x} f_i'(x) \quad .$$

Then $f_i(x) = \int_a^x f_i'(t)dt = \int_a^x g_i(t)dt$ and

$$f(x) = \int_a^x (g_1(t) - g_2(t)dt = \int_0^x g(t)dt$$

where $g(t) = g_1(t) - g_2(t)$ is in $BV[a,b]$ since each $g_i(x)$ is monotone. Conversely, if $f(x) = \int_a^x g(t)dt$ with $g \in BV[a,b]$ then $g = m_1 - m_2$ with m_1, m_2 monotone. It suffices to show that

$$f_i(x) = \int_a^t m_i(t)dt$$

is convex. Suppose, to this end, that $x_2 \geq x_1$. Then

$$f_i(x_2) - f_i(x_i) = \int_{x_1}^{x_2} m_i(t)dt \geq m_i(x_1)(x_2-x_1) \quad .$$

Similarly, if $x_2 \leq x_1$, then $m(\delta) \leq m(x_1)$ for $x_2 \leq \delta \leq x_1$ and

$$f_i(x_2) - f_i(x_1) = -\int_{x_2}^{x_1} m_i(t)dt \geq m_i(x_1)(x_2-x_1) \ .$$

Thus f_i has a tangent at every point and so is convex ([4]).
Thus $f(x) = f_1(x) - f_2(x) \in K(\Omega)$. □

In R^n it appears more difficult to describe $K(\Omega)$. The following direct result gives some additional information to Theorem 1. Let $P(R^n)$ denote the real valued polynomials in n variables.

Theorem 3: $P(R^n) \subset K(R^n)$.

Proof: (i) Let $n = 1$. It suffices to note that each x^n is in $K(R)$. Since both $\max(x^n,0)$ and $-\min(x^n,0)$ are convex functions for any n, $x^n \in K(R)$.

(ii) Let $n = 2$. Molluzzo has shown that any polynomial in two variables can be written as a sum of polynomials in linear combinations of the variables. That is

$$(1) \qquad P(x,y) = \sum_{i=1}^{r} P_i(a_i x + b_i y)$$

where each P_i is a polynomial of a single variable and the coefficients a_i and b_i are real.

Now each P_i in (1) can be decomposed into convex parts and since a convex function of a linear function is convex it follows that

$$P(x,y) = \sum_{i=1}^{m} P_i^{+}(a_i x + b_i y) - \sum_{i=1}^{m} P_i^{-}(a_i x + b_i y)$$

is the desired decomposition where $P_i = P_i^{+} - P_i^{-}$ is the convex decomposition of P_i .

(iii) Now suppose that for $n \leq n_1$ and any polynomial in n variables

$$(2) \qquad P(x_1,\ldots,x_n) = \sum_{i=1}^{m} P_i(\ell_i(x_1,\ldots,x_n))$$

where the ℓ_i are linear functionals in R^n and the P_j are polynomials of a single variable. We show by induction that (2) holds for n_1+1, and since the same decomposition can be used as in (1), the result is established.

It suffices to consider

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} x_i^{n_i} = \hat{P}(x_1,\ldots,x_{n-1})x_n^{n_n} \text{ . Now}$$

$$\hat{P}(x_1,\ldots,x_{n-1}) = \sum_{i=1}^{m} \hat{P}_i(L_i(x_1,\ldots,x_{n-1})) \text{ so that}$$

$$P(x_1,\ldots,x_n) = \sum_{i=1}^{m} \hat{P}_i(w_i)x_n^{n_n} \text{ with } w_i = L_i(x_1,\ldots,x_{n-1}).$$

Now by (2), $\hat{P}_i(w_i)x_n^{n_n}$ can be written as

This page from an earlier version shows explicitly how to decompose in two dimensions.

Let $k + \ell = m$. Consider solving

$$\sum_{i=0}^{m} (1+i)^r d_i = \delta_{k,r}/\binom{m}{r} \qquad 0 \le r \le m$$

for real d_i . This can be done since $(C_{ir}) = ((i+1)^r)$ is a Vandermond matrix ([5]) with distinct entries. Now let $\bar{b}_i = (d_i)^{1/m}$ and $\bar{a}_i = (i+1)\bar{b}_i$. Then

$$\sum_{i=0}^{m} \bar{a}_i^r \bar{b}_i^{m-r} = \sum_{i=0}^{m} (i+1)^r d_i^{r/m} d_i^{1-r/m} = \delta_{k,r}/\binom{m}{r} ,$$

and

$$\sum_{i=0}^{m} d_i((1+i)x+y)^m = \sum_{i=0}^{m} (\bar{a}_i x + \bar{b}_i y)^m = \sum_{i=0}^{m} \sum_{j=0}^{m} \bar{a}_i^j \bar{b}_i^{m-j} \binom{m}{j} x^j y^{m-j}$$

$$= \sum_{j=0}^{m} \frac{\binom{m}{j}\delta_{k,j}}{\binom{m}{j}} x^j y^{m-j}$$

$$= \binom{m}{k}/\binom{m}{k} x^k y^\ell = x^k y^\ell .$$

It follows that any polynomial in two variables has a decomposition as in (1). Now each P_i in (1) can be decomposed into convex parts and since a convex function of a linear function is convex it follows that

$$P(x,y) = \sum_{i=1}^{m} P_i^+(a_i x + b_i y) - \sum_{i=1}^{m} P_i^- (a_i x + b_i y)$$

is the desired decomposition where $P_i = P_i^+ - P_i^-$ is the convex decomposition of P_i .

(iii) Now suppose that for $n \le n_0$ and any polynomial in n variables

(2) $$P(x_1,\ldots,x_n) = \sum_{i=1}^{m} P_i(L_i(x_1,\ldots,x_n))$$

where the L_i are linear functionals in R^n and the P_i are polynomials of a single variable. We show by induction that (2) holds for n_{0+1} , and since the same decomposition can be used as in (1), the result is established.

$$\hat{P}_i(w_i)(x_n^n) = \sum_{j=1}^{m'} P_{j,i}(a_{ij}w_i + b_j x_n)$$

$$= \sum_{j=1}^{m'} P_{j,i}(a_{ij}x_1 + \ldots + a_{nj}x_n)$$

and one has the desired decomposition. □

Remarks (i) It is an immediate corollary of this theorem and the Stone-Weierstrass theorem that $K(R^n)$ is dense in $C(R^n)$ with the uniform topology, so that Theorem 1 follows in this case.

(ii) In the case $n = 2$ the decomposition in (1) is due to Molluzzo [2] with a proof which also gives information on the number of polynomials necessary in (1). Note that this information will not extend to $n \geq 2$.

(iii) While Theorem 3 does not characterize $K(R^n)$ we know from Theorem 2 that $K(R^n)$ does not coincide with certain other classes. In particular there are differentiable functions (such as $f(x) = \int_0^x t \sin \frac{1}{t}\, dt$) which are not in $K[-1,1]$ while the Cantor function is monotone without even being absolutely continuous. Thus $K(R^n)$ does not absorb either Lipshitz or differentiable functions on R^n.

In the case of quadratic functions on a complex Hilbert space, H, one has the following type of convex quadratic decomposition:

Theorem 4: Suppose A is a continuous linear mapping H. Then $g(x) = (Ax,x)$ can be written as

$$g(x) = g_1(x) + ig_2(x)$$

where each $g_i(x)$ is the difference of convex quadratic functions.

Proof: Let $A = B_1 + iB_2$ with B_1, B_2 self adjoint.

Let $B_i = B_i^+ - B_i^-$ where

$$B_i^+ = \frac{B_i+\sqrt{B_i^2}}{2}\ ,\ B_i^- = \frac{-B_i+\sqrt{B_i^2}}{2}\ .$$

Note that the square root of B_i^2 exists since B_i^2 is self adjoint and positive ([1]). It follows that $B_i^+B_i^- = 0$ and that B_i^+, B_i^- are positive ([1]). Thus

$$g(x) = (B_1^+x,x) - (B_1^-x,x) + i((B_2^+x,x) - (B_2^-x,x))$$

which is the desired decomposition since, for a bounded operator T, one has that $(T\,x,x)$ is convex when T is self adjoint positive. □

References

(1) G. Bachmann and L. Naraci, Functional Analysis, Academic Press, (1965).

(2) J.C. Molluzzo, A Representation Theorem for Polynomials of Two Variables, Am. Math. Monthly, 82, (4), (1975), 385-387.

(3) L. Nachbin, Elements of Approximation Theory, D. Van Nostrand, (1967).

(4) R.T. Rockafellar, Convex Analysis, Princeton, (1967).

(5) R.R. Stoll and E.J. Wong, Linear Algebra, Academic Press, (1968).

(6) A.E. Taylor, General Theory of Functions and Integration, Blaisdell Publishing Co., (1965).