

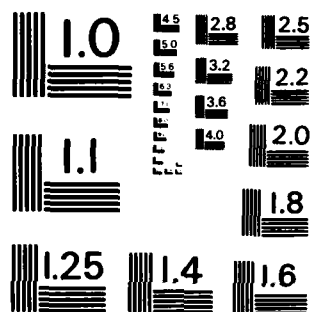
AD-A145 428 AN ALGORITHM FOR LEAST SQUARES PROJECTIONS ONTO THE
INTERSECTION OF SHIFT..(U) IOWA UNIV IOWA CITY DEPT OF
STATISTICS AND ACTUARIAL SCIENCE.. R L DYKSTRA JAN 84
UNCLASSIFIED TR-103 N00014-83-K-0249 F/G 12/1

1/1

NL

END
DATE
FILMED

10 84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

13

AD-A145 428

An Algorithm for Least Squares Projections
onto the Intersection of Shifted, Convex Cones

by

Richard L. Dykstra
Department of Statistics
and Actuarial Science
University of Iowa

DTIC
ELECTE
SEP 12 1984

B

Technical Report #103
January 1984

This work was partially supported by ONR Contract N000 14-83-K-0249.

DISTRIBUTION STATEMENT A

Approved for public release
Distribution Unlimited

84 09 06 055

DTIC FILE COPY

- A -

ABSTRACT

A commonly occurring problem is that of minimizing least squares expressions subject to restrictions on the solution. Dykstra (1983) has given a simple algorithm for solving these types of problems when the constraint region can be expressed as a finite intersection of closed, convex cones. Here it is shown that this algorithm must still work correctly even when each cone is allowed to be arbitrarily translated (as long as the intersection is non-empty). This allows the algorithm to be applied to a much larger collection of problems than previously indicated.

Key words and phrases: Least squares projections, regression, cones constraints, translations, minimizations, dual cones, iterations.

AMS Classification numbers: Primary: 49099
: 65015

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By PER CALL TO	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



An Algorithm for Least Squares Projections
onto the Intersection of Shifted Convex Cones

Richard L. Dykstra

1. Introduction.

The problem of obtaining least squares projections subject to various constraints is a frequently occurring problem in many areas. For example, the area of isotonic regression usually is concerned with obtaining least squares vectors which must satisfy certain partial order restrictions (see Barlow, Bartholomew, Bremner and Brunk (1972) and Robertson and Wright (1981)). Another example concerns finding the closest (least squares) convex (concave) function through a set of points in the plane (see Hildreth (1954) and Wu (1982)).

Many times, the constraint region can be written as a finite intersection of simpler constraint regions. This raises the possibility of using iterative schemes based upon the projections onto the simpler regions for solving the overall problem.

John Von Neumann (1950) has shown that if the constraint region is an intersection of two subspaces, cyclic iterative projections onto the individual subspaces must converge to the desired projection. Norbert Wiener (1955) independently proved a version of this theorem in a slightly different setting.

Dykstra and Robertson (1982) have developed an iterative procedure for finding projections of rectangular arrays onto the class of arrays with nondecreasing rows and columns based only upon one-dimensional smoothings. Later, Dykstra (1983) extended this approach to the general framework of projections onto the intersection of closed convex cones. This procedure is based upon finding projections onto the individual cones, and reduces to Von Neumann's and Weiner's method when the cones are also subspaces.

It is the purpose of this paper to show that Dykstra's (1983) algorithm can be extended to work for projections onto a finite non-empty intersection of shifted (translated) closed, convex cones. In particular, this means that least squares projections under general linear constraints of the form $\sum_{i=1}^n a_i x_i \leq (=) b$ can be handled by Dykstra's algorithm even when $b \neq 0$. This clearly follows by writing $\{x: \sum_{i=1}^n a_i x_i \leq b\}$ as

$$\{x: \sum_{i=1}^n a_i x_i \leq 0\} - (-b/a_1, 0, \dots, 0).$$

Constraints of the form

$$y_i \leq x_i \leq z_i, \quad i = 1, \dots, n$$

where y and z are fixed vectors fall into the translated cone framework by writing

$$\{x: y_i \leq x_i \leq z_i \quad \forall_i\}$$

as

$$(\{x: x_i \geq 0\} - (-y_1, -y_2, \dots, -y_n)) \cap (\{x: x_i \leq 0\} - (-z_1, -z_2, \dots, -z_n)).$$

Other examples of cone constraints given in Dykstra (1983) can be generalized to translated cone constraints.

2. Notation and Setting.

We denote n -dimensional real coordinate space by R^n , and let g and $w(w_i > 0)$ be fixed points in R^n . The inner product of x and y (with respect to w) is given by

$$(x, y) = \sum_{i=1}^n x_i y_i w_i.$$

The corresponding inner product norm of x is defined as

$$(2.1) \quad \|x\| = (x, x)^{\frac{1}{2}} = \left(\sum_{i=1}^n x_i^2 w_i \right)^{\frac{1}{2}}.$$

A closed (in the metric) subset K of R^n is a closed convex cone if $x, y \in K$; $a, b \geq 0$ implies $ax + by \in K$. The dual cone of a closed, convex cone K is defined as

$$(2.2) \quad K^* = \{y \in R^n; (y, x) = \sum_{i=1}^n y_i x_i w_i \leq 0 \quad \forall x \in K\}.$$

Of course K^* is also a closed, convex cone with the property that $K^{**} = K$.

A commonly occurring problem is to find the x which will

$$(2.3) \quad \underset{x \in C}{\text{Minimize}} \quad \|g - x\|$$

where C is a closed, convex set ($x, y \in C$ implies $\alpha x + (1-\alpha)y \in C$ $\forall \alpha \in [0,1]$). A vector $g^* \in C$ achieves the minimal value in (2.3) iff

$$(2.4) \quad (g - g^*, g^* - f) \geq 0 \quad \text{for all } f \in C.$$

If C is actually a closed, convex cone, we may replace (2.4) by

$$(2.5) \quad \begin{array}{ll} \text{i)} & (g - g^*, g^*) = 0, \quad \text{and} \\ \text{ii)} & (g - g^*, f) \leq 0 \quad \text{for all } f \in C. \end{array}$$

(See Theorem 7.8 of Barlow et al. (1972).)

Note that (2.5) implies that if C is a closed, convex cone and g^* solves (2.3), then $g - g^* \in C^*$.

3. The Algorithm.

We wish to consider problems of the form

$$(3.1) \quad \underset{x \in \bigcap_{i=1}^p (K_i - b_i)}{\text{Minimize}} \quad \|g - x\|$$

where K_i is a closed, convex cone in R^n , $b_i \in R^n$,
 $K_i - b_i = \{x - b_i; x \in K_i\}$, and $\bigcap_{i=1}^r (K_i - b_i) \neq \emptyset$. We assume that we can
 find the vector in $K_i - b_i$ which will

$$(3.2) \quad \begin{array}{l} \text{Minimize} \quad \|f - x\| \\ x \in K_i - b_i \end{array}$$

for any f and any i , and wish to use this ability to solve (3.1).

We note that if $P(f|C)$ denotes the projection of $f \in R^n$
 onto the closed, convex set C , then for a closed, convex cone K

$$(3.3) \quad P(f|K_i - b_i) = P(f + b_i|K_i) - b_i.$$

Thus

$$(3.4) \quad f - P(f|K_i - b_i) = (f + b_i) - P(f + b_i|K_i) \in K_i^* \quad \text{for all } f$$

by (2.5) ii).

We shall make extensive use of (3.4).

Our proposed algorithm is identical to that given in Dykstra (1983)
 except that we allow projections onto shifted, closed, convex cones.
 Our scheme can be succinctly stated with the aid of the following
 notation:

- i) For any positive integer n , we define $n(\bmod r) = i$ if
 $n = kr + i$ for integers k and i where $1 \leq i \leq r$.
- ii) Initially, set $n = 0$, $g_0 = g$, and $I_i = 0 \in R^n$, $i = 1, \dots, r$.

The iterative procedure is to

- 1) Set $g_{n+1} = P(g_n - I_{n(\text{mod } r)} | K_{n(\text{mod } r)} - b_{n(\text{mod } r)})$, and
 then update $I_{n(\text{mod } r)}$ by resetting it equal to
 (3.5) $g_{n+1} - (g_n - I_{n(\text{mod } r)})$.
 2) Replace n by $n+1$ and go to 1).

We refer the reader to Dykstra (1983) for further elaboration on the algorithm and its uses. This procedure requires only the ability to find projections onto the K_i (see (3.3)). These individual projections are often easy to program and quick to execute, and hence can be combined to solve rather difficult optimization problems.

In particular, the algorithm applies to quadratic programming problem with a finite number of constraints of the form $\sum_{i=1}^n a_i x_i \leq b$.

4. Proof of the Algorithm.

To simplify the proof for $n \geq 1$, we will write g_n in (3.5) as $g_{k,i}$ when $n = (k-1)r + i$, $1 \leq i \leq r$. (In other words, $g_{k,i}$ is the projection onto the i^{th} shifted cone during the k^{th} cycle.)

In similar fashion we depict

$$(4.1) \quad I_{k,i} = \begin{cases} g_{k,i} - (g_{k,i-1} - I_{k-1,i}), & \text{if } 2 \leq i \leq r \\ g_{k,1} - (g_{k-1,r} - I_{k-1,1}), & \text{if } i = 1. \end{cases}$$

We will also have need of the following lemma, a proof of which is found in Dykstra (1983).

Lemma 4.1. Suppose a sequence of nonnegative real numbers $\{a_n\}_{n=1}^{\infty}$ is such that $\sum_{n=1}^{\infty} a_n^2 < \infty$. Then there exists a subsequence $\{a_{n_j}\}_{j=1}^{\infty}$ such that

$$\sum_{m=1}^{n_j} a_m a_{n_j} \rightarrow 0 \text{ as } j \rightarrow \infty.$$

We now establish the fundamental result of the paper.

Theorem 4.1. The vectors g_n defined in (3.5) converge to the true solution, say g^* , of the problem defined in (3.1) as $n \rightarrow \infty$.

Proof. Since $\bigcap_{i=1}^r (K_i - b_i) \neq \emptyset$, we may assume WLOG that $0 \in \bigcap_{i=1}^r (K_i - b_i)$. Then $b_i \in K_i$ and the true solution g^* exists uniquely. Note from (4.1) that

$$g_{n,i-1} - g_{n,i} = I_{n-1,i} - I_{n,i}, \quad i = 2, \dots, r, \text{ and} \quad (4.2)$$

$$g_{n-1,r} - g_{n,1} = I_{n-1,1} - I_{n,1}.$$

Thus, in general $(I_{0,i} = 0)$, for $i \geq 2$

$$\begin{aligned} \|g_{n,i-1} - g^*\|^2 &= \|(g_{n,i} - g^*) + (I_{n-1,i} - I_{n,i})\|^2 \\ (4.3) \quad &= \|g_{n,i} - g^*\|^2 + \|I_{n-1,i} - I_{n,i}\|^2 \\ &\quad + 2(g_{n,i} + b_i, I_{n-1,i} - I_{n,i}) - 2(g^* + b_i, I_{n-1,i} - I_{n,i}). \end{aligned}$$

Note that $(g_{n,i} + b_i, I_{n,i}) = 0$ (by (2.5)).

Moreover, since $g_{n,i} + b_i \in K_i$ and $-I_{n-1,i} \in K_i^*$ (by (3.4)), the next to last term is nonnegative.

Similarly

$$\begin{aligned}
 \|g_{n-1,r} - g^*\|^2 &\geq \|g_{n,1} - g^*\|^2 + \|I_{n-1,1} - I_{n,1}\|^2 \\
 (4.4) \qquad &\quad - 2(g^* + b_1, I_{n-1,1} - I_{n,1}).
 \end{aligned}$$

Repeated application of (4.3) and (4.4) together with addition and the telescoping property of the last term yields

$$\begin{aligned}
 \|g - g^*\|^2 &\geq \|g_{n,r} - g^*\|^2 + \sum_{k=1}^n \sum_{\ell=1}^r \|I_{k-1,\ell} - I_{k,\ell}\|^2 \\
 (4.5) \qquad &\quad + 2 \sum_{i=1}^r (g^* + b_i, I_{n,i}) \text{ for all } n.
 \end{aligned}$$

Since the last sum is nonnegative ($g^* + b_i \in K_i$, $-I_{n,i} \in K_i^*$), we know

$$(4.6) \quad \sum_{k=1}^{\infty} \sum_{\ell=1}^r \|I_{k-1,\ell} - I_{k,\ell}\|^2 < \infty,$$

and hence

$$\begin{aligned}
 \|I_{n-1,\ell} - I_{n,\ell}\| &= \|g_{n,\ell-1} - g_{n,\ell}\| \quad (\ell \geq 2) \text{ and} \\
 (4.7) \qquad &\quad \|I_{n-1,1} - I_{n,1}\| = \|g_{n-1,r} - g_{n,1}\| \rightarrow 0 \text{ as } n \rightarrow \infty.
 \end{aligned}$$

We note that (4.5) implies that $g_{n,r}$ and $g - g_{n,r} = I_{n,1} + \dots + I_{n,r}$ are uniformly bounded.

Next we show, that there exists a subsequence $\{n_j\}$ such that

$$(4.8) \quad \limsup_j (I_{n_j,1} + I_{n_j,2} + \dots + I_{n_j,r}, g_{n_j,1} - f) \leq 0 \quad \forall f \in \bigcap_1^r (K_i - b_i)$$

To see this, note that

$$(4.9) \quad \begin{aligned} & (I_{n,1} + \dots + I_{n,r}, g_{n,1} - f) \\ &= \sum_{i=1}^r (I_{n,i}, g_{n,1} - f - (g_{n,i} + b_i)) \quad (\text{since } (I_{n,i}, g_{n,i} + b_i) = 0) \\ &= \sum_{i=2}^r (I_{n,i}, g_{n,1} - g_{n,i}) + \sum_{i=1}^r (-I_{n,i}, f + b_i). \end{aligned}$$

The last sum is nonpositive since $f + b_i \in K_i$ and $-I_{n,i} \in K_i^*$.

For the first part, we use the Cauchy-Schwarz Inequality to say

$$\begin{aligned} & \left| \sum_{i=2}^r (I_{n,i}, g_{n,1} - g_{n,i}) \right| \\ & \leq \sum_{i=2}^r \|I_{n,i}\| \|g_{n,1} - g_{n,i}\| \\ & \leq \sum_{i=2}^r \left(\sum_{m=1}^n \|I_{m,i} - I_{m-1,i}\| \right) \left(\sum_{\ell=2}^r \|g_{n,\ell-1} - g_{n,\ell}\| \right) \\ & = \left(\sum_{m=1}^n \sum_{i=2}^r \|I_{m,i} - I_{m-1,i}\| \right) \left(\sum_{\ell=2}^r \|g_{n,\ell-1} - g_{n,\ell}\| \right) \\ & = \sum_{m=1}^n a_m a_n \end{aligned}$$

where $a_n = \sum_{\ell=2}^r \|g_{n,\ell-1} - g_{n,\ell}\| = \sum_{\ell=2}^r \|I_{n,\ell} - I_{n-1,\ell}\|$ (see (4.7)).

Since

$$a_n^2 \leq 2^{r-2} \sum_{\ell=2}^r \|I_{n,\ell} - I_{n-1,\ell}\|^2,$$

(4.6) implies that $\sum_{n=1}^{\infty} a_n^2 < \infty$. Thus lemma (4.1) can be employed to yield (4.8). Moreover, since the $g_{n,r}$ are uniformly bounded, we may assume that we have chosen a subsequence such that (4.8) holds and

$g_{n_j,r}$ converges, say to h . Note that (4.7) insures that $g_{n_j,\ell}$ also converges to h for every ℓ , and hence that

$h \in \bigcap_{i=1}^r (K_i - b_i)$ (the K_i are closed). In addition, since the

$g_{n_j,r} - g = I_{n_j,1} + \dots + I_{n_j,r}$ are uniformly bounded and converge to $h-g$, we may use (2.4) to argue $h = g^*$.

Finally, in a manner similar to (4.5), we can show

$$\begin{aligned} \|g_{n_j,r} - g^*\|^2 &= \|g_{n,r} - g^*\|^2 + \sum_{m=j+1}^n \sum_{\ell=1}^r \|I_{m,\ell} - I_{m-1,\ell}\|^2 \\ (4.10) \quad &+ 2 \sum_{m=n_j+1}^n \sum_{\ell=1}^r ((g_{m,\ell} + b_\ell) - (g^* + b_\ell), I_{m-1,\ell} - I_{m,\ell}). \end{aligned}$$

We may write the last sum as

$$(4.11) \quad 2 \sum_{m=n_j+1}^n \sum_{\ell=1}^r (g_{m,\ell} + b_\ell, I_{m-1,\ell})$$

$$(4.12) \quad -2 \sum_{m=n_j+1}^n \sum_{\ell=1}^r (g_{m,\ell} + b_\ell, I_{m,\ell})$$

$$(4.13) \quad + 2 \sum_{\ell=1}^r (g^* + b_\ell, I_{n,\ell})$$

$$(4.14) \quad - 2 \sum_{\ell=1}^r (g^* + b_\ell, I_{n_j,\ell}).$$

Each term of (4.11) is nonnegative ($g_{m,\ell} + b_\ell \in K_\ell$, $-I_{m-1,\ell} \in K_\ell^*$).

Each term of (4.12) is zero by (2.5). Each term of (4.13) is nonnegative for the same reason as (4.11).

Finally, we can write (4.14) as

$$\begin{aligned} & -2[(g^*, \sum_{\ell=1}^r I_{n_j,\ell}) + \sum_{\ell=1}^r (b_\ell, I_{n_j,\ell})] \\ &= -2[(\lim_{j \rightarrow \infty} g_{n_j,1}, \sum_{\ell=1}^r I_{n_j,\ell}) + \sum_{\ell=1}^r (b_\ell, I_{n_j,\ell})] \\ &\rightarrow -2[\lim_{j \rightarrow \infty} [(g_{n_j,1}, \sum_{\ell=1}^r I_{n_j,\ell}) + \sum_{\ell=1}^r (b_\ell, I_{n_j,\ell})]] \\ &= -2 \lim_{j \rightarrow \infty} [\sum_{\ell=2}^r (I_{n_j,\ell}, g_{n_j,1} - g_{n_j,\ell})] \end{aligned}$$

by taking f to be Q in 4.9. This clearly equals zero by the way $\{n_j\}_{j=1}^\infty$ was chosen. Then noting that the left side of (4.10) goes to zero as $j \rightarrow \infty$, it easily follows that $g_{n,r} \rightarrow g^*$ as $n \rightarrow \infty$. This clearly is good enough by (4.7).

REFERENCES

1. Barlow, R.E., Bartholomew, D.J., Bremner, J.M., and Brunk, H.D. (1972). Statistical Inference Under Order Restrictions, New York: John Wiley.
2. Dykstra, Richard L. (1983). "An Algorithm for Restricted Least Squares Regression," Journal of the American Statistical Association, 78, 837-842.
3. Dykstra, Richard L. and Robertson, Tim (1982). "An Algorithm for Isotonic Regression for Two or More Independent Variables," Annals of Statistics, 10, 708-716.
4. Hildreth, C. (1954). "Point Estimates of Ordinates of Concave Functions," Journal of the American Statistical Association, 49, 598-619.
5. Robertson, Tim and Wright, F.T. (1981). "Likelihood Ratio Test For and Against a Stochastic Ordering Between Multinomial Populations," Annals of Statistics, 9, 1248-1257.
6. Von Neumann, John (1950). Functional Operators (Vol. II). Princeton, N.J.: Princeton University Press.
7. Wiener, N. (1955). "On Factorization of Matrices," Comment Mathematici Helvetici, 29, 97-111.
8. Wu, Chein-Fu (1982). "Some Algorithms for Concave Regression and Isotonic Regression," Management Science (forthcoming).

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 103	2. GOVT ACCESSION NO. AD-A145 428	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Algorithm for Least Squares Projections onto the Intersection of Shifted, Convex Cones		5. TYPE OF REPORT & PERIOD COVERED Technical - August 1984
		6. PERFORMING ORG. REPORT NUMBER 103
7. AUTHOR(s) Richard L. Dykstra		8. CONTRACT OR GRANT NUMBER(s) N00014-83-K-0249
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics and Actuarial Science The University of Iowa Iowa City, Iowa 52242		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 N. Quincy Arlington, VA 22217		12. DATE JAN 1984
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Least squares projections, regression, cones, constraints, translations, minimizations, dual cones, iterations.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A commonly occurring problem is that of minimizing least squares expressions subject to restrictions on the solution. Dykstra (1983) has given a simple algorithm for solving these types of problems when the constraint region can be expressed as a finite intersection of closed, convex cones. Here it is shown that this algorithm must still work correctly even when each cone is allowed to be arbitrarily translated (as long as the intersection is nonempty). This allows the algorithm to be applied to a much larger collection of problems than previously indicated.		

DD FORM 1473

EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

