

January 2010

Mathematical Themes in Economics, Machine Learning, and Bioinformatics

Matt Bogard

Western Kentucky University, matt.bogard@wku.edu

Follow this and additional works at: http://digitalcommons.wku.edu/econ_fac_pub



Part of the [Agricultural and Resource Economics Commons](#), and the [Applied Statistics Commons](#)

Recommended Repository Citation

Bogard, Matt, "Mathematical Themes in Economics, Machine Learning, and Bioinformatics" (2010). *Economics Faculty Publications*. Paper 4.

http://digitalcommons.wku.edu/econ_fac_pub/4

This Other is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Economics Faculty Publications by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.

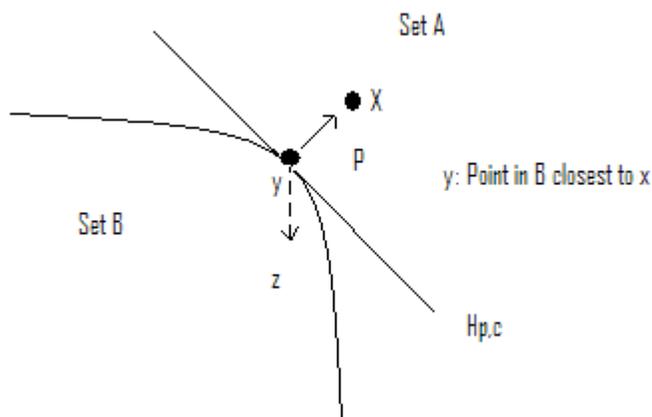
Mathematical Themes in Economics, Machine Learning, and Bioinformatics

Graduate students in economics are often introduced to some very useful mathematical tools that many outside the discipline may not associate with training in economics. This essay looks at some of these tools and concepts, including constrained optimization, separating hyperplanes, supporting hyperplanes, and 'duality.' Applications of these tools are explored including topics from machine learning and bioinformatics.

Most all economics students are introduced to the concept of the utility maximization problem (UMP), where a utility function, say $U(x,y)$ is introduced and maximized subject to some constraint such as the budget constraint say $I - P_x \cdot X - P_y \cdot Y = 0$. They are also introduced to the expenditure minimization problem (EMP) where total expenditures represented by say $E = P_x \cdot X + P_y \cdot Y$ are minimized subject to a utility constraint, say $U = U(x,y)$. By use of the lagrangian multiplier method it turns out that the optimal vector in the UMP is the same for the EMP. In other words, UMP and EMP are 'dual' problems. The demand function is derived from the UMP and the expenditure function is derived from the EMP. Typically for first semester graduate students, the second order conditions for UMP and EMP are presented in terms of the bordered Hessian matrix. Second order conditions imply that for a maximum in a constrained optimization problem, the hessian must be negative definite, and it must be positive definite for a constrained minimization problem.

Later in graduate school, if economics students use Mas-Colell, Whinston, and Green's *Microeconomic Theory*, they will be introduced to a different concept of duality. Green's mathematical appendix defines two important theorems as follows:

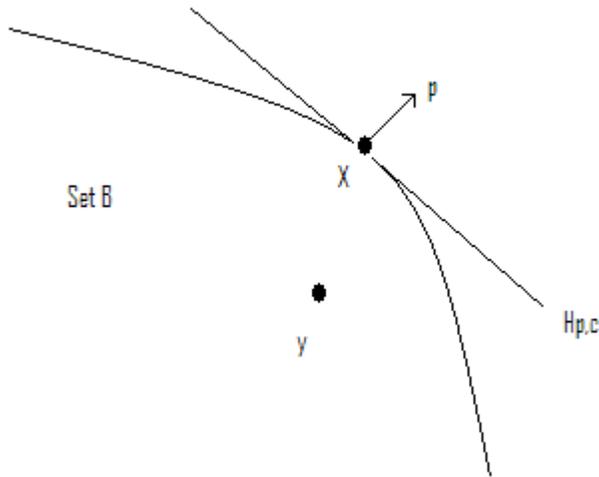
Separating Hyperplane Theorem:



(Image adapted from Green, 1995)

- Suppose $A, B \subset \mathbb{R}^N$ and $A \cap B = \emptyset$. Suppose B is a convex set and closed, and that $x \notin B$. Then there is a $p \in \mathbb{R}^N$ with $p \neq 0$, and a value $c \in \mathbb{R}^N$ such that $p \cdot x > c$ and $p \cdot y < c$ for every $y \in B$. i.e. There is a hyperplane that separates A and B leaving A and B on opposite sides. (Green, 1995)

Supporting Hyperplane Theorem:

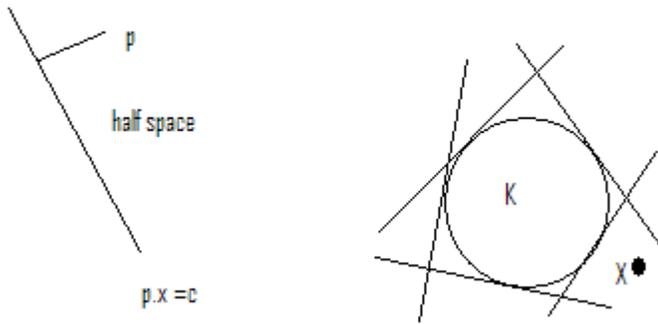


(Image adapted from Green, 1995)

- Suppose $B \subset \mathbb{R}^N$ is convex and that x is not an element of the interior set of B (i.e. $x \notin \text{Int } B$). Then there is $p \in \mathbb{R}^N$ with $p \neq 0$ such that $p \cdot x \geq p \cdot y$ for every $y \in B$. i.e. the hyperplane $H_{p,c}$ supports the set B . (Green, 1995)

APPLICATION: Consumer Choice Theory from Economics

As Mas-Colell states, a consequence of the separating hyperplane theorem is that if K is a convex set, then there is a half space (defined by a normal vector and a hyperplane) containing K and excluding x .

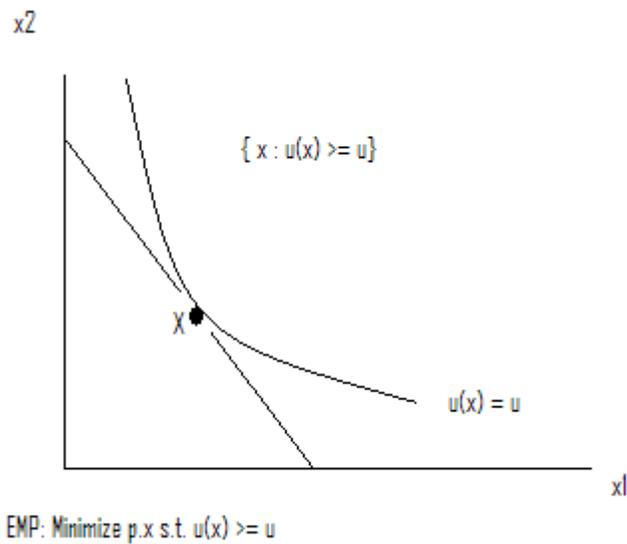


(Image adapted from Green, 1995)

i.e. a closed convex set can be 'dually' defined by the sum of its half spaces. A support function is then defined:

- A support function for K can be defined as $\mu_K(p) = \inf\{p \cdot x : x \in K\}$ and $K = \{x \in \mathbb{R}^n : p \cdot x \geq \mu_K(p) \text{ for every } p\}$

It turns out that the expenditure function $e(p,u)$ derived from the expenditure minimization problem in microeconomics is the support function for the set $\{x : u(x) \geq u\}$



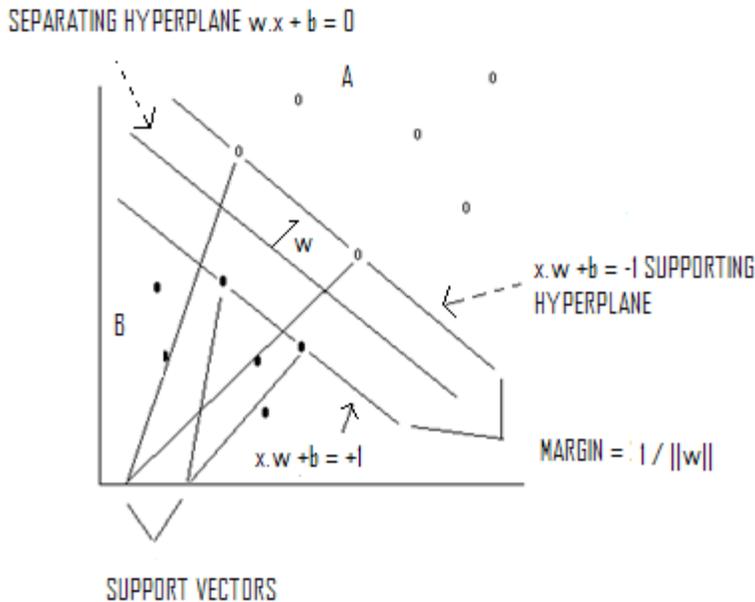
APPLICATION : Support Vector Machines

'SVM's select the hyperplane that maximizes the margin of separation between two classes from among all separating hyperplanes. The SVM classifies inputs into two classes using a hyperplane in high dimensional space' (Wu,2006). Ultimately, this machine learning or data mining classifier makes use of the same concepts from mathematical economics that many economics graduate students may be familiar with, such as supporting hyperplanes and separating hyperplanes, as well as constrained optimization. According to Fletcher, a SVM involves deriving w and b and using them to describe the training data $\{x,y\}$ as such:

$$x_i \cdot w + b \geq +1 \text{ for } y_i = +1 \text{ and } x_i \cdot w + b \leq -1 \text{ for } y_i = -1 \text{ or}$$

$$y_i (x_i \cdot w + b) - 1 \geq 0 \forall i$$

Margin is defined using vector geometry to be $1 / \|w\|$ where w is a vector normal to a separating hyperplane that is equidistant from the supporting hyperplanes (described by support vectors) that support the different classes or sets of observations in the training data.



Maximizing margin becomes an optimization problem involving the use of the Lagrangian multiplier method familiar to economics students. In setting up the Lagrangian, the term $\alpha^T H \alpha$ is formulated where $H_{ij} = y_i y_j x_i x_j$. In this context, the 'linear kernel' is described as a dot product, being $k(x_i, x_j) = x_i^T x_j$. If data is not linearly separable, then alternative kernel functions are used to map non-linearly separable data into a feature space where the data are linearly separable. Examples may include the use of radial

basis functions, polynomial, and sigmoidal kernel functions. The kernel matrix must be positive semi-definite.

APPLICATION: Bioinformatics

In 'Comparing Kernals for Predicting Protein Binding Sites from Amino Acid Sequence' Wu, et al recognize that kernel functions in SVM's allow you to capitalize on domain specific knowledge. They compare 3 kernal types including the identity kernel, sequence alignment kernel, and an amino acid sequence substitution matrix kernel for predicting protein-protein, protein-DNA, and Protein-RNA binding sites. They are careful to ensure that their kernel formulations satisfy the Mercer conditions which require kernel matrices to be positive semi definite.

References:

Mas-Colell, Winston and Green. Microeconomic Theory. Oxford University Press. (1995)

Wu, F., Olson, B., Dobbs, D., and Honavar, V. "Comparing Kernels For Predicting Protein Binding Sites From Amino Acid Sequence". *IEEE Joint Conference on Neural Networks, Vancouver, Canada*, IEEE Press. Vol. In press, 2006.

Tristen Fletcher. Support Vector Machines Explained. March 1,2009.