On the universal method to solve extremal problems.

Mathematical and economic applications.

by

Jan Brinkhuis,

Erasmus University Rotterdam,

January 7, 2005

For Vladimir Mikhailovich Tikhomirov on his seventieth birthday.

Abstract. Some applications of the theory of extremal problems to mathematics and economics are made more accessible to non-experts.

1. The following fundamental results are known to all users of mathematical techniques, such as economists, econometricians, engineers and ecologists: the fundamental theorem of algebra, the Lagrange multiplier rule, the implicit function theorem, separation theorems for convex sets, orthogonal diagonalization of symmetric matrices. However, full explanations, including rigorous proofs, are only given in relatively advanced courses for mathematicians. Here, we offer short and easy proofs. We show that all these results can be reduced to the task of solving a suitable extremal problem. Then we solve each of the resulting problems by a universal strategy.

2. The following three practical results, each earning their discoverers the Nobel prize for Economics, are known to all economists and econometricians: Nash bargaining, the formula of Black and Scholes for the price of options and the models of Prescott and Kydland on the value of commitment. However, the great value of such applications of the theory of extremal problems deserves to be more generally appreciated. The great impact of these results on real life examples is explained. This, rather than mathematical depth, is the correct criterion for assessing their value.

Econometric Institute Report EI 2005-02

1 Introduction

Plan. The theory of extremal problems, also called optimization problems, arose from the scientific curiosity of people who admired the chain of beautiful pearls consisting of individual extremal problems that were completely solved analytically—that is, by a formula—by great people who came up with a brilliant idea to solve the problem at hand. With hindsight, we can now say that—almost—all these problems are in some sense standard. There is one universal strategy to solve all these problems (cf [Ioffe-Tikhomirov], [Alekseev-Tikhomirov-Fomin], [Tikhomirov]). This realization should not diminish their beauty and power in our eyes. In this short fragment, a few convincing examples of problems from two of the main sources are given: mathematics and economics. More information is given in [Brinkhuis-Tikhomirov]. The paper has been made self-contained for the benefit of students, and experts in other fields than optimization.

Acknowledgements. I would like to express my thanks to Jan Boone and Shuzhong Zhang for help with the economic applications. I would like to extend my thanks to Grażyna and Samantha Brinkhuis for helpful comments. I am grateful to Jan van de Craats for making the Figure. Finally, part of the research for this paper was done while I enjoyed the hospitality of the Chinese University of Hong Kong.

2 Universal solution strategy: four step method

All—or at least almost all—problems of optimization that can be solved analytically, can be solved by means of one method, which is based on the Weierstrass theorem. This method can be presented as a four step method. We will only consider problems to *minimize* a function f; maximization problems can be written as minimization problems by means of a minus sign: maximizing $f(x)$ is equivalent to minimizing $-f(x)$.

Four step method:

1. Give a formalization and establish existence of a solution (= minimum) $\widehat{x}$ by means of the Weierstrass theorem.
2. Write necessary conditions. Give an admissible variation x_α, $\alpha \geq 0$, of $\widehat{x}$, calculate $g(\alpha) = f(x_\alpha)$, and write the condition $g(\alpha) \geq g(0)$.
3. Analysis of this condition.
4. Conclusion.

On step 1. The Weierstrass theorem states that a continuous function on a nonempty closed and bounded domain assumes its global minimum at some point. If the domain is not bounded but the objective function f is *coercive*—that is, $f(x) \to \infty$ for $|x| \to \infty$—then we are also assured of existence of a solution, as adding the constraint $|x| \leq C$ for a sufficiently large constant C does not change the solvability status of the problem, and reduces it to a problem where the Weierstrass theorem can be applied.

3 Mathematical applications.

The first source of extremal problems is the tendency of humans to reach 'the essence of everything'. This has led to many *mathematical* results that can be obtained using optimization methods. Here the yardstick to measure the quality of applications of optimization methods is the depth of the results. Here we ask how *difficult* it would be to establish these results by other methods. If this is difficult, then we hold this application in a high esteem and consider it convincing. This class of results will be illustrated by the following examples.

1. Fundamental theorem of algebra;
2. Orthogonal diagonalization of symmetric matrices;
3. Main theorem of classical analysis (tangent space theorem, or, equivalently, implicit function theorem or inverse function theorem);
4. Main theorem of convex analysis (separation theorem for point outside closed convex set, or, equivalently, other separation theorems, supporting hyperplane theorem, existence of subgradients or nontriviality of dual cone);
5. Lagrange multiplier rule.

All these results are widely used. They are known to each user of mathematical methods. However, their proofs are usually only given in relatively advanced courses. Our aim is to make the point that *all these results can be proved easily by one universal strategy.*

4 Fundamental theorem of algebra

Theorem 1. *Each nonconstant polynomial* $p(x) = a_0 \ldots + a_n x^n$ *of degree* $n \geq 1$ *has a real or complex root.*

Proof.

1. $f(x) = |p(x)| \to \min$, where x runs over the complex plane $\mathbb{C}$, has a solution $\widehat{x}$ as f is coercive. To prove the theorem, it suffices to show that the value of this problem is zero.

   Assume $\widehat{x} = 0$: this can be done without loss of generality to the argument (replace x by $x - \widehat{x}$). Then it remains to show that $a_0 = f(0)$ is zero.

2. Let u be a solution of $a_0 + a_k x^k = 0$ where k is the smallest positive number for which the coefficient a_k is nonzero. That is, u is one of the k-th roots of $-a_0/a_k$.

   Define $x_\alpha = \alpha u$ for $\alpha > 0$ and $g(\alpha) = f(x_\alpha)$.

Then $g(\alpha) \geq g(0) \Rightarrow$
$|a_0 + a_k(\alpha u)^k + o(\alpha^k)| \geq |a_0|$.

3. $|(1-\alpha^k)a_0 + o(\alpha^k)| \geq |a_0| \Rightarrow a_0 = 0$.
4. The theorem holds true.

Q.E.D.

Remark 1. The idea of this proof is due to d'Alembert. The presentation of the proof given above emphasizes the beautiful essence of it.

Remark 2. This is one of the rare examples of an optimization problem of interest that requires the necessary conditions of order greater than two (you need the necessary conditions of order k for local minimality of the function $g(\alpha),\ \alpha \geq 0$ at $\alpha = 0:$ as $g^{(i)}(0) = 0,\ 1 \leq i \leq k-1$, the following k-th order necessary condition holds true,

$$g^{(k)}(0) = -k!|a_0| \geq 0,$$

and this gives $a_0 = 0$ as required).

Remark 3. Note in passing that this analysis may serve as a motivation for the use of complex numbers and as an illustration of the golden words of Hadamard that the best way to arrive at a truth about the real numbers is by means of the complex numbers. To this end, the theorem should be formulated without using the complex numbers:

every polynomial of one variable with real coefficients can be written as a product of linear and quadratic factors with real coefficients.

5 Main property symmetric matrices

Theorem 2. *Each real-valued symmetric $n \times n$-matrix A can be turned into a diagonal matrix D by conjugation with a suitable orthogonal matrix X:*

$$X^T AX = D,\ X^T X = I.$$

The sense of this result is that each quadratic form can be brought into diagonal form

$$y \mapsto d_1 y_1^2 + \cdots + d_n y_n^2$$

by an orthogonal transformation of the variables.

Define $h(M)$ to be the sum of the squares of the off-diagonal elements of a given square matrix M. This quantity is always nonnegative, and it is precisely zero if the matrix is a diagonal matrix.

We will use the following notation: for each $n \times n$-matrix M we split up both the rows and the columns into two groups, the first two and the last $n-2$; this gives the following block decomposition $M = \begin{bmatrix} M_{ll} & M_{lr} \\ M_{rl} & M_{rr} \end{bmatrix}$.

Proof.

1. The problem $f(X) = h(X^T AX) \to \min$, $X^T X = I$ has a solution $\widehat{X}$ by Weierstrass. To prove the theorem, it suffices to prove that the value of this problem is zero. We assume $\widehat{X} = I$ without restriction of the generality of the argument (replace X by $\widehat{X}X$ and, correspondingly, $\widehat{X}^T A\widehat{X}$ by A). Therefore, to prove the theorem, it suffices to prove that a symmetric matrix A is a diagonal matrix if $h(X^T AX) \geq h(A)$ for all orthogonal matrices X.

2. Define
$$\tilde{X} = \begin{bmatrix} P & 0_{2\times n-2} \\ 0_{n-2\times 2} & I_{n-2} \end{bmatrix}$$
where P is chosen such that $P^T A_{ll} P$ is diagonal and $P^T P = I_2$; as A_{ll} is a symmetric 2×2-matrix, this can be done by giving an explicit formula for P.

   We write the condition
$$f(\tilde{X}) - f(I) \geq 0$$
and calculate the left hand side of this inequality: this gives $-2a_{12}^2$ and so we get $-a_{12}^2 \geq 0$. Indeed, note to begin with that
$$\tilde{X}^T A\tilde{X} = \begin{bmatrix} P^T A_{ll} P & P^T A_{lr} \\ A_{rl} P & A_{rr} \end{bmatrix},$$

   Now use that each column of A_{lr} has the same euclidian length, as the corresponding column of $P^T A_{lr}$, as P is orthogonal. This shows that the sum of the squares of all off diagonal entries except the $(1,2)$ and $(2,1)$-entries is the same for A as for $\tilde{X}^T A\tilde{X}$. Therefore, $f(\tilde{X}) - f(I) = 0 - 2a_{12}^2$, as required.

3. It follows from $-a_{12}^2 \geq 0$ that $a_{12} = 0$. In the same way one can show that $a_{ij} = 0$ for all $1 \leq i, j \leq n$ with $i \neq j$. That is, A is a diagonal matrix.

4. The statement of the theorem holds true.

Q.E.D.

Remark 4. This proof has some novel features: it proceeds by the solution of one optimization problem by means of the four step method. The proof by optimization that is given in textbooks uses the Lagrange multiplier rule and proceeds by solving a sequence of optimization problems.

Remark 5. The proof above is algorithmic in spirit. In fact it is related to an efficient class of algorithms, the Jacobi methods.

6 Main theorem of classical analysis

Theorem 3. Tangent space theorem. *Let* $\bar{x} \in \mathbb{R}^n,\ h \in C^1(\bar{x}, \mathbb{R}^m)$*, and* $h(\bar{x}) = 0, \quad h'(\bar{x})(\mathbb{R}^n) = \mathbb{R}^m$*. Then the system of equations* $h(x) = 0$ *defines in the neighborhood of* $\bar{x}$ *a manifold of codimension* m *with tangent space the kernel of* $h'(\bar{x})$*, in the following sense:*

there is $\varepsilon > 0$ *such that for each solution* $v \in B_n(0, \varepsilon)$ *of the linear system* $h'(\bar{x})x = 0$*, there exists a unique vector* $r(v) \in B_m(0, \varepsilon)$ *with*

$$h(\bar{x} + v + h'(\bar{x})^T r(v)) = 0,\ r(v) = o(v).$$

For the geometric interpretation, note that the vectors $h'(\bar{x})^T v,\ v \in \mathbb{R}^n$ are precisely the vectors that are orthogonal to the solution set of the linear system $h'(\bar{x})x = 0$.

We will only sketch the proof, as this allows to focus on the essential ideas without being distracted by some details that are routine but technical.

Proof. We assume $\bar{x} = 0$, as we may without restriction to the generality of the argument. We choose $\varepsilon > 0$ sufficiently small. In the proof we will use repeatedly that $h'(x) \approx h'(0)$ if $x \approx 0$.

1. $f(r) = |h(v + h'(0)^T r)| \to \min,\ |r| \le \varepsilon$ has a unique solution $r(v)$ by Weierstrass.

2. Fermat: $f'(r) = 0 \Rightarrow$

$$f'(r) = |h(v + h'(0)^T r)|^{-1} h(v + h'(0)^T r) h'(v + h'(0)^T r) h'(0)^T = 0.$$

3. **There are no stationary points.** This follows by using

   - $h'(x) \approx h'(0)$ if $x \approx 0$,
   - the matrix $h'(0)h'(0)^T$ is invertible by the assumption

$$h'(0)(\mathbb{R}^n) = \mathbb{R}^m.$$

There are no boundary points of minimum. For a boundary point $\bar{r}$ one has

$$f(\bar{r}) \approx |h(0) + h'(0)(v + h'(0)^T)\bar{r}| =$$

$$|h'(0)h'(0)^T \bar{r}| \ge \varepsilon \|h'(0)h'(0)^T\|^{-1},$$

whereas $f(0) = |h(v)| \approx |h(0) + h'(0)v| = 0$. Therefore, $f(\bar{r}) > f(0)$.

Point of minimum is non-smooth. It follows that $r(v)$ is a non-smooth point of f, and so $h(v + h'(0)^T r(v)) = 0$. This shows as well that the value of the problem is zero.

Point of minimum is unique. For two points of minimum r_1 and r_2, one has

$$0 = h(v + h'(0)^T r_1) - h(v + h'(0)^T r_2) \approx h'(0)h'(0)^T (r_1 - r_2)$$

and so $r_1 = r_2$.

Point of minimum is small.

Take linear approximation at origin:

$$\langle h'(0)^T, v + h'(0)^T r(v)\rangle + o(v) = 0 \Rightarrow r(v) = o(v).$$

4. The statement of the theorem holds true.

Q.E.D.

Remark 6. The well-known implicit and inverse function theorems are variants of the tangent theorem and follow immediately from it.

7 Main theorem of convex analysis

Theorem 4. *A nonempty closed convex set C can be separated from each point $p \notin C$ by a hyperplane.*

Proof.

1. $f(x) = |x - p|^2 \to \min,\ x \in C$ has a solution $\widehat{x}$ by Weierstrass.
2. Take an arbitrary $c \in C$.

   Write $v = \widehat{x} - p$ and $r = c - \widehat{x}$.

   Define $x_\alpha = \widehat{x} + \alpha r$ for $\alpha > 0$ and $g(\alpha) = f(x_\alpha)$.

   Then $g(\alpha) \geq g(0)$ for α sufficiently small, by the minimality of $\widehat{x}$,

   $\Rightarrow |v + \alpha r|^2 \geq |v|^2$.
3. Expanding and simplifying gives $\langle v, r\rangle \geq 0$.
4. The hyperplane $\langle v, x - \widehat{x}\rangle = 0$ separates the set C and the point p.

Q.E.D.

Remark 7. The well-known first and second separation theorems, the supporting hyperplane theorem, the existence of subgradients for convex functions, and the nontriviality of dual cones are variants of this result and follow immediately from it.

8 Lagrange multiplier rule

For constrained extremal problems, the four step method can be simplified: in step 2 one should apply immediately the appropriate necessary conditions (the Lagrange multiplier rule or the Karush-Kuhn-Tucker theorem) instead of choosing a variation. This is no essential difference as the proofs of these necessary conditions proceed by giving a suitable variation. Now we illustrate this by proving the Lagrange multiplier rule by means of the four step method.

Theorem 5. Lagrange multiplier rule. *Let $\widehat{x} \in \mathbb{R}^n,\ f \in C^1(\widehat{x}),\ F \in C^1(\widehat{x}, \mathbb{R}^m)$ with $F(\widehat{x}) = 0$ be given. If $\widehat{x}$ is a point of local minimum of the problem*

$$f(x) \to \min,\ F(x) = 0,$$

then there exist Lagrange multipliers $\lambda_0 \in \mathbb{R}$ and $\lambda \in (\mathbb{R}^m)'$, not both zero, such that the Lagrange equations hold true:

$$\lambda_0 f'(\widehat{x}) + \lambda F'(\widehat{x}) = 0.$$

Proof.

1. We consider the problem $f(x) \to \min,\ F(x) = 0$.
2. We assume that $F'(\widehat{x})$ has rank m, otherwise the Lagrange equations hold clearly, with $\lambda_0 = 0$. Choose an arbitrary $v \in \ker F'(\widehat{x})$. We choose an admissible variation $x_\alpha,\ \alpha \geq 0$, with $x_\alpha = \widehat{x} + \alpha v + o(\alpha)$, as we may by virtue of the tangent space theorem. Define $g(\alpha) = f(x_\alpha)$. Write $g(\alpha) \geq g(0)$.
3. We get $g'(0) = 0$, that is, $f'(\widehat{x})v = 0$. Therefore, as v is an arbitrary element of $\ker F'(\widehat{x})$, it follows, using the theory of linear equations, that $f'(\widehat{x})$ is a linear combination of the rows of $F'(\widehat{x})$.
4. The theorem holds true.

Q.E.D.

9 Economic applications.

The second source of optimization problems is the pragmatic wish to understand the world and to act in it in the best possible way. As a result there are many *economic* results that can be derived using optimization methods. Here the name of the game is totally different.

To quote the famous words of the great economist Marshall: *"A good mathematical theorem dealing with economic hypotheses is very unlikely to be good economics"*. The success of an economic application of optimization methods is measured by the light it sheds on examples that are important in real life. It is not the mathematical depth of the results that counts here. The greater the impact of the result on our understanding of economic questions, the higher the esteem in which we hold these applications of optimization methods. We will illustrate this class by three applications that have had a great impact on our understanding of the world and that therefore earned their discoverers the Nobel prize in Economics.

1. Bargaining (Nash): what is a fair bargain?
2. Options (Black and Scholes): what is the right price for options?
3. Time (in)consistency (Prescott and Kydland): the value of commitment.

Our aim is to make the following point. *Just as everybody is subject to the laws of gravity of Newton whether he is aware of these laws or not, so everybody is confronted with the consequences of these three applications.* This will be made explicit below.

In the next sections we give some convincing economic applications. In each case we give the simplest mathematical model that illustrates the economic phenomenon at hand, and we will emphasize the impact of the model. Here we will not display the details of the analysis with the four step method.

10 Nash bargaining

Impact. In the seminal paper [Nash], John Nash considered the problem what would be a fair outcome if two people have an opportunity to bargain. This is a question of great interest as most of us are daily confronted with the necessity to negotiate. The solution proposed by Nash is convincing for the following three reasons.

1. It is intuitive: each player gets his outside option plus a part of the netto cake (that is, total surplus minus player's outside options) and the part you get depends on your bargaining power;
2. the axioms of Nash are very convincing;
3. there is a non-cooperative bargaining game that has the Nash outcome as solution. The game proceeds as follows. You and I have to divide a cake. First you make a proposal how to divide it, $(x, 1-x)$. Then I have the choice to accept it or to reject it. If I reject, then I can make in the next period a proposal to divide it, $(y, 1-y)$. Then you have the choice to accept or to reject. This game has in principle an infinite horizon. However, the problem is that in each period a part of the cake disappears. The reason for this can be that the quality of the cake deteriorates; just by discounting, or because there is a chance that one of us will leave the game and then the cake disappears. The solution of the game is the Nash outcome of the bargaining opportunity. We will present the solution of

Nash, but for the underlying bargaining game we refer to [Fudenberg-Tirole]. We only note that to solve the game (with its recursive structure) the Bellman equations are used. The bargaining power (exogenous in the model of Nash) now depends on how risk averse or how impatient you are (for example, my discount factor δ can be higher than yours).

We illustrate bargaining with a numerical example.

Example. Jennifer and Britney are in a position to barter goods but have no money with which to facilitate exchange.

Jennifer's goods	*Utility to Jennifer*	*Utility to Britney*
ball	1	4
Britney's goods		
bat	2	1
box	2	2

What should they do?

The analysis of Nash, to be given below, leads to the following solution: Jennifer gives her ball to Britney, Britney gives her bat to Jennifer, and a coin is tossed: if it is heads, then Britney gives her box to Jennifer (otherwise she keeps it).

Now we are going to describe the result of Nash. A *bargaining opportunity* is modelled as a convex set F in the plane $\mathbb{R}^2$ together with a point $v \in F$, such that the set of all $x \in F$ with $x \geq v$ is closed and bounded. It is assumed that there exists an element $w \in F$ for which $w > v$. The two people can either come to an agreement—the points of F represent the possible agreements—or not: if not, then they will have to accept as outcome the point v. The point v is called *the disagreement point* or *the outside option*. For each outcome, the utility for person i is the i-th coordinate of the point corresponding to the outcome ($i = 1, 2$).

We illustrate these concepts for the numerical example above.

Example. The bargaining opportunity can be modelled by the set F of linear combinations $p_1(4, -1) + p_2(-1, 2) + p_3(-2, 2)$ with coefficients in $[0, 1]$ and with outside option v the origin $(0, 0)$. The solution given above corresponds to the following choice of coefficients: $p_1 = 1,\ p_2 = 1,\ p_3 = \frac{1}{2}$.

Fair outcome rule. To address the question what is a fair outcome of the bargain, we have to consider the more general problem of finding a fair rule ϕ to assign to an arbitrary bargaining opportunity (F, ϕ) an outcome $\phi(F, v)$ in F. This rule should only depend on the underlying preferences of the two individuals and not on the utility functions that are used to describe these preferences (if a utility function u is used to describe a preference, then the utility functions that describe the same preference are the affine transformations of u: $cu + d$ with $c > 0$). It is natural to impose the following axioms on this rule. We need the following concept: a pair (F, v) is called *symmetric* if there exist affine transformations of the two utility functions such that after transformation the following implication holds : if $(a, b) \in F$, then $(b, a) \in F$, that is, the set F is symmetric with respect to the line $x_1 = x_2$; this means that both persons are in a symmetrical situation, as they are confronted with exactly the same bargaining opportunity.

1. There are no $x \in F$ for which $x > \phi(F, v)$.

   This represents the idea that no Pareto improvements are possible. We emphasize that it does not really represent individual maximization. Take for example a monopolist who maximizes his own profit, but not total welfare. That is, there is an allocation where both the monopolist and the consumer are better off. The first axiom states that such a situation does not occur. All advantages of trade are used. The solution has the property that there is no allocation where everyone is better off.

2. If (E, v) is another bargaining opportunity with $E \subseteq F$, and $\phi(F, v) \in E$, then $\phi(E, v) = \phi(F, v)$.

   This represents the idea that eliminating from F feasible points (other than the disagreement point) that would not have been chosen should not effect the solution.

3. If (F, v) is symmetric and utility functions are chosen such that this is displayed, then $\phi(F, v)$ is a point of the form (a, a).

   This represents equality of bargaining skill: if the two people have the same bargaining position, then they will get the same utility out of it.

Theorem 7 (**Nash**). *There is precisely one rule ϕ that satisfies the three axioms above: $\phi(F, v)$ is the unique solution of the following optimization problem:*

$$f(x) = (x_1 - v_1)(x_2 - v_2) \to \max, \ x \in F, \ x \geq v.$$

Proof. For each pair (F, v) the problem has a solution by Weierstrass, and this solution is unique as taking minus the logarithm of the objective function, we can transform the problem into an equivalent convex problem with a strictly convex objective function

$$-\ln(x_1 - v_1) - \ln(x_2 - v_2).$$

Now we can assume, by an appropriate choice of the utility functions of the two individuals, that the two coordinates of this solution are equal to 1 and that $v_1 = v_2 = 0$. Then $(1, 1)$ is the solution of the problem $x_1 x_2 \to \max$, $x \in F$, $x \geq 0_2$. The set F is contained in the halfplane G, given by $x_1 + x_2 \leq 2$; let us argue by contradiction: assume that there exists a point $x \in F$ outside this halfplane, then the segment with endpoints v and x is contained in F, by the convexity of F, and this segment contains points near v with $x_1x_2 > 1$, which contradicts the maximality of $(1,1)$. The axioms above give immediately that $\phi(G, 0_2) = (1,1)$ as G is symmetric, and so that $\phi(F, 0_2) = (1,1)$ as $\phi(G, 0_2) \in F$. *Q.E.D.*

Remark 8. In particular, to find the Nash solution for any given bargaining opportunity, a convex optimization problem has to be solved. It is recommended to verify this for the numerical example given above.

11 Pricing of options by Black and Scholes

Impact. In the seminal paper [Black-Scholes], Black and Scholes gave a method for pricing options. This work revolutionized the practice of finance. Options play an essential role in the economy, as they allow firms to protect themselves against risks that they cannot bear on their own. Options already exist for thousands of years, but how to price them always seemed a matter of taste, depending on one's prediction of the future development of the price of the stock, and moreover, on one's willingness to take risks. A surprising insight of the work of Black and Scholes was that the value of an option is exactly the same to everyone. The spectacular development of the option markets would not have been possible without the work of Black and Scholes.

Simplified model. The pricing method of Black and Scholes involves the solution of a Bellman equation in continuous time (this is a partial differential equation). In this section, we present a later—simpler—discrete time model: the case of one (short) period and two scenarios. This simplified model involves three assets—stock, bond and option— and two scenarios—stock goes up or stock goes down.

- **Stock.** A stock has current price p; in the first scenario its value will go down to $v^{(1)}$, in the second scenario it will go up to $v^{(2)}$.
- **Bond.** In addition there is a bond, a riskless asset; we put its current price at 1 and we can achieve, by a normalization, that its value will remain 1.
- **Option.** In the third place, there is a European call option on the stock, with exercise price w. This is the right to buy the stock after the completion of the investment period for the price w.
- **Value option after period.** The option has after the period value $(v^{(i)} - w)_+$ in scenario i (here we use the notation $a_+ = a$ if $a > 0$ and $a_+ = 0$ otherwise). Indeed, if $v^{(i)}$, the value of the stock after the period, turns out to be higher than w, then the owner of the option can buy the stock for w and sell it for $v^{(i)}$, thus making a profit $v^{(i)} - w$. Otherwise he cannot make any profitable use of it.
- **Probability of scenarios.** Nothing is known about the probability of the scenarios.

- **Pricing of options problem.** The problem is for which price p_o the option should be sold.

- **Absence of arbitrage opportunities.** The arbitrage opportunities that we consider can be defined, loosely speaking, as investment strategies, where we get money now, but never have to pay anything. This idea can be turned into the following precise definition. An arbitrage opportunity is a vector (x_1, x_2, x_3) for which the following inequalities hold true:

  - $px_1 + x_2 + p_o x_3 < 0$,
  - $v^{(1)}x_1 + x_2 + (v^{(1)} - w)_+ x_3 \geq 0$,
  - $v^{(2)}x_1 + x_2 + (v^{(2)} - w)_+ x_3 \geq 0$.

  Here, x_1 is the number of stocks bought, x_2 is the number of bonds bought, and x_3 is the number of options bought. The numbers x_1, x_2, x_3 can be positive, zero or negative ('going short'). Therefore, the first inequality means that you get money if you accept to hold this investment. The second (resp. third) investment means that you will not have to pay any money after the period in scenario 1 (resp. 2).

- **Formula of Black and Scholes**. The analysis of Black and Scholes shows that the price of the option is determined uniquely by the absence of arbitrage opportunities, and is given by the following formula:

$$p_o = (v^{(2)} - v^{(1)})^{-1}[(v^{(2)} - p)(v^{(1)} - w)_+ + (p - v^{(1)})(v^{(2)} - w)_+]. \qquad (*)$$

We illustrate the formula of Black and Scholes with a numerical example.

Example. We consider some stock and some bond, each of which are worth 100 euro now. The value of the stock after one month will be either 200 euro or 50 euro; the probabilities of these two scenarios are not necessarily equal, in fact they are not known. The value of the bond after one month will always be still 100 euro. Now we consider an option that gives the right to buy the stock after one month for 110 euro, whatever its value will be.

How should this option be priced?

The formula of Black and Scholes gives 30 euro as the correct price for this option..

Now we are going to explain how the formula of Black and Scholes can be derived. We will give another way to present the formula; this has the advantage that it allows a concrete probabilistic interpretation.

For a full understanding, a more general context is considered: n assets and m scenarios. Afterwards, we will apply it to three assets—stock, bond, option— and two scenarios—'stock goes up' and 'stock goes down'.

Consider the following set-up:

- m assets and one investment period
- $p_1, \ldots, p_m$ prices at beginning investment period
- $v_1, \ldots, v_m$ values at end period
- $x_1, \ldots, x_m$ investments at beginning (here short-selling, that is, negative x_i, are allowed)
- cost investment $p^T \cdot x$
- final value investment $v^T \cdot x$

Now we model risk as well:

- we consider n scenarios
- let $v_1^{(j)}, \ldots, v_m^{(j)}$ be the values in scenario j

We assume that the $v_i^{(j)}$ are known, but not the probabilities of these scenarios. It might be surprising at first sight that the prices $p_1, \ldots, p_m$ are not arbitrary. The reason for this is that it is reasonable to assume that there are no *arbitrage opportunities*. Arbitrage opportunities can be defined, loosely speaking, as investment strategies, where we get money now, but never have to pay anything. This idea can be turned into the following precise definition. An arbitrage opportunity is a vector x for which

$$p^T \cdot x < 0, \quad V^T x \geq 0,$$

where V is the $m \times n$-matrix $(v_i^{(j)})_{ij}$.

Theorem 8. *If there are no arbitrage opportunities, then there exists $y \geq 0$ with $p = Vy$.*
Proof.

1. $f(x) = p^T \cdot x \to \min,\ V^T x \geq 0$ is solvable by the assumption that there are no arbitrage opportunities.
2. Karush-Kuhn-Tucker: there exists $y \geq 0$ with $p = Vy$

Q.E.D.

Remark 9. A risk-less asset can be defined to be one for which the value in each scenario equals the price at the beginning of the investment period, such as a bond. If the collection of n assets contains a risk-less asset, then y can be interpreted as a probability distribution on the collection of scenarios: indeed, $y_1 + \cdots + y_m = 1,\ y_i \geq 0,\ 1 \leq i \leq m$. Then the other equations of the system $p = Vy$ have the following probabilistic interpretation: for each asset i the price

p_i at the beginning of the period equals the expectation of its value at the end of the period with respect to the probability distribution y:

$$p_i = y_1 v_i^{(1)} + \cdots + y_m v_i^{(m)}.$$

Remark 10. For example, if the prices at the beginning of the period of all but one asset are known, then the result above gives information about the price at the beginning of the period of the remaining asset: it lies between the extremal values of the following optimization problem (with given $m \times n$-matrix V and with given prices $p_1, \ldots, p_{n-1}$:

$$f(p, y) = p_n \rightarrow \text{extr},\ p = Vy,\ y \geq 0.$$

In particular, if the extremal values of this problem coincide, then the price of this remaining asset is determined. Now we will consider such a problem and this will give the pricing method for options of Black and Scholes.

Theorem 9. Formula of Black and Scholes: a fair price for options. *Consider a stock with current price p and with value $v^{(1)}, v^{(2)}$ in the two possible scenario's. Assume that in the first scenario the value goes down and in the second scenario it goes up: $v^{(1)} < p < v^{(2)}$. In addition, there is a bond, a risk-less asset, with current price 1 and with value 1 in each of the two possible scenario's. We consider the problem of pricing a European call option on a stock with exercise price w. Show that the absence of type A arbitrage determines the price p_o of the option. It is given by the following formula of Black and Scholes:*

$$p_o = y_1(v^{(1)} - w)_+ + y_2(v^{(2)} - w)_+,$$

where y_1, y_2 are determined by

$$y_1, y_2 \geq 0,\ y_1 + y_2 = 1,$$

and

$$p = y_1 v^{(1)} + y_2 v^{(2)}.$$

This theorem gives a linear system of three equations and three unknowns: y_1, y_2 and p_o. Solving this system, we get the formula $(*)$ above.

The formula of Black and Scholes allows the following interpretation:

there exists a unique probability distribution y_1, y_2 for the two possible scenarios for which the price of the stock and the bond are equal to their expected value after the investment period.

Proof. Theorem 8 with three assets (stock, bond and option) and two scenario's, leads to the following system of three linear equations in the unknowns $p_o,\ y_1,\ y_2$:

$$p = y_1 v^{(1)} + y_2 v^{(2)},$$

$$1 = y_1 + y_2,$$

$$p_o = y_1 (v^{(1)} - w)_+ + y_2 (v^{(2)} - w)_+.$$

This is seen to have a unique solution, and this solution has y_1 and y_2 nonnegative. *Q.E.D.*

12 Time (in)consistency

Impact. In the seminal paper [Kydland-Prescott], Kydland and Prescott have made clear that in order for the policies of central banks to be successful, these banks should be able to commit themselves to carrying out these policies irrespective of wishes of politicians. Their work has played a decisive role in the creation of independent central banks. This is believed to have led to a long period with low inflation.

The model of Kydland and Prescott. The model is illustrated in Figure 1. Now we

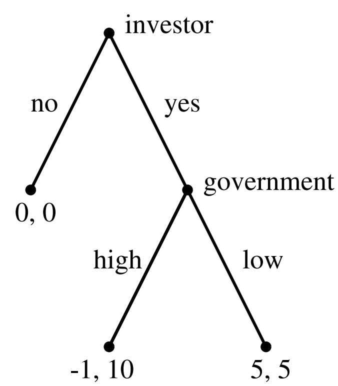

Figure 1: Pay off structure.

give a verbal description of this model. At $t = 0$ an investor decides whether or not to invest in a country. At $t = 1$ the government decides on the tax rate on capital. We assume the pay off structure is as follows. If the investor does not invest in the country, both the investor and the government get pay offs normalized to zero. If the investor does invest in the country and the government implements a low tax rate, the investor and the government both get a pay off equal to 5. If, however, the investor invests and the government at $t = 1$ chooses a high tax rate the investor loses and gets -1 while the government gains a big tax receipt and gets pay off equal to 10.

If the government can commit at $t = 0$ to the action it takes at $t = 1$ it would commit to a low tax rate. Commitment can be achieved by making a law stating that the tax rate is low (that is, changing the law leads to a cost bigger than 5, making the switch to a high tax rate unprofitable later on). Or the government may try to get a reputation for low tax rates. In case of a government trying to fight inflation, the commitment comes from creating a central bank that is independent.

If the government cannot commit, it will choose its optimal plan at each moment it has to take a decision. That implies that once the investor has invested, the government chooses the high tax rate since a pay off of 10 exceeds a pay off of 5. The investor (using backward induction) sees this happening and chooses not to invest because a pay off of 0 exceeds a pay off of -1.

Clearly, the commitment solution is never worse than the no commitment solution (because you can always commit to the Bellman solution) and in the case above it is clearly better. That is, commitment has value. An example often used here is Ulysses tied to the mast to listen to the Sirens. In that way, he committed not to go to them while he could still listen.

Also note that in a deterministic world commitment is always better than no commitment. In a world with uncertainty, there is trade off. With commitment one can not respond to unexpected events. In the example above, you may have committed to a low tax rate, but the country may end up in fiscal crisis and you cannot respond by increasing the tax rate.

Relation with optimization. One of the ideas of Kydland and Prescott can be explained in terms of the difference between solutions between dynamic optimization problems by the following two different methods: Pontryagin's Maximum Principle and the Bellman equation. The former gives an *open loop solution*: this specifies all decisions to be taken during the planning period (in technical terms: a control function $u(t)$ of time t). The latter gives a *closed loop solution* (or *feedback solution*): this solution gives a contingency plan, that is, a strategy how to take decisions during the planning period depending on how the situation will be in future (in technical terms: a control function $u(t, x)$ of time t and the current state x). At first sight it might seem that a closed loop solution is always preferable, because of its greater flexibility. However, this is not always the case, as the model from [Kydland-Prescott] that is given above, demonstrates. If a decision maker can make the commitment to carry out a Pontryagin solution, he is in a strong position. A decision maker who uses a Bellman solution, is vulnerable to manipulations.

Specifically, in the model above, the commitment solution ('to choose at $t = 1$ a low tax rate') corresponds to the open loop solution or the Pontryagin solution to a dynamic optimization problem; the no commitment solution ('to choose at $t = 1$ a high tax rate') corresponds to the Bellman solution or closed loop solution: at each moment in time you choose your best action.

References

V.M. Alekseev, V.M. Tikhomirov, S.V. Fomin, *Optimal Control*, Consultants Bureau, New York (1987).

F. Black, M. Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 81 (May-June 1973), 637-59.

J. Brinkhuis, V.M. Tikhomirov, *Optimization: insights and applications; lunch, dinner, dessert*, Princeton University Press, to appear August 2004.

D. Fudenberg, J. Tirole *Game Theory*, Cambridge, MA: MIT Press, (1991).

A.D. Ioffe and V.M. Tikhomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam-New York-Oxford (1979).

F.E. Kydland, E.C. Prescott, *Rules rather than discretion: the inconsistency of optimal plans*, Journal of Political Economy 85, nr 3 June (1977), 473-492.

J.F. Nash, Jr, *The bargaining problem*, Econometrica 18 (1950) 155-162.

V.M. Tikhomirov, *Stories about Maxima and Minima*, Mathematical World, volume 1, American Mathematical Society (1990).