12^{th} International Workshop
for Young Mathematicians
Probability Theory and Statistics
Kraków, 20-26 September 2009
pp. 43 - 51

APPLICATIONS OF THE KANTOROVICH-RUBINSTEIN MAXIMUM PRINCIPLE IN THE THEORY OF MARKOV OPERATORS

HENRYK GACKI

The author dedicates the walk to the memory of
Professor Andrzej Lasota
his Master and long time Friend.

Abstract. The main purpose of this note is to show a new sufficient condition for the asymptotic stability of Markov operators acting on the space of signed measures. This condition is applied to stochastically perturbed dynamical systems, and iterated function system (IFS). They are formulated in terms of adjoint operators. This approach simplifies further applications.

Our results are based on two principles. The first one is the LaSalle invariance principle used in the theory of dynamical systems. The second is related to the Kantorovich-Rubinstein theorems concerning the properties of probability metrics.

1. Introduction

The Kantorovich-Rubinstein maximum principle is stimulated by *the Kantorovich-Rubinstein duality theorem* ([8, 12, 14]). The duality theorem has a long and colourful history, apparently originating in 18th century (1781) work of Monge on *the transport of mass problem* ([11]). Let μ_1 and μ_2 are two Borel probability measures given on a separable metric space (X, ϱ) and let V be the space of probability measures on $X \times X$ such that its projections on the first and second coordinates are μ_1 and μ_2, respectively i.e.

$$V = V(\mu_1, \mu_2) := \big\{ b \in \mathcal{M}_1(X \times X);\ \Pi_1 b = \mu_1,\ \Pi_2 b = \mu_2 \big\}$$

The measures μ_1 and μ_2 may be viewed as the initial and final distribution of mass and V as the space of admissible transference plan. If the unit cost of shipment from x to y is $\varrho(x, y)$ then the total cost is

$$\int\limits_{X \times X} \varrho(x, y) b(dx, dy).$$

Thus we see that minimization of transportation costs can be formulated in terms of finding a distribution on $X \times X$ whose marginals are fixed, and such that the double integral of the cost function is minimal. The duality Kantorovich-Rubinstein theorem can be stated as follows

$$\inf_{b\in V} \int\limits_{X\times X} \varrho(x,y)b(dx,dy) = \sup_{f\in\mathcal{H}} \int\limits_{X} f(x)(\mu_1-\mu_2)(dx).$$

During the study of the classical Monge transport problem L. V. Kantorovich and G. S. Rubinstein discovered some interesting properties of the functional

$$\varphi_\mu(f) = \int\limits_X f(x)\mu(dx) \qquad \text{for} \qquad \mu=\mu_1-\mu_2.$$

In particular this functional always admits its maximum value on the set of Lipschitzean functions with Lipschitz constant $L \leq 1$. Moreover, every function which realizes the maximum of φ_μ satisfies the condition

$$|f(x)-f(y)| = \varrho(x,y),$$

for some $x, y \in X, x \neq y$. This property will be called *the Kantorovich-Rubinstein maximum principle.* The maximum principle will be used to prove a new sufficient condition for asymptotic stability for Markov operators acting on the space of signed measures. This condition generalizes some previous results of A. Lasota ([9]), K. oskot and R. Rudnicki ([10]).

2. Kantorovich-Wasserstein metric in the space of measures

Let (X,ϱ) be a Polish space, i.e., a separable, complete metric space. By $\mathcal{B}_X$ we denote the σ–algebra of Borel subsets of X and by $\mathcal{M}$ the family of all finite (nonnegative) Borel measures on X. By $\mathcal{M}_1$ we denote the subset of $\mathcal{M}$ such that $\mu(X)=1$ for $\mu \in \mathcal{M}_1$. The elements of $\mathcal{M}_1$ will be called *distributions.* Further let

$$\mathcal{M}_{sig} = \{\mu_1-\mu_2 :\ \mu_1,\mu_2 \in \mathcal{M}\},$$

be the space of finite signed measures.

Let c be a fixed element of X. For every real number $\alpha \geq 1$ we define the sets $\mathcal{M}_{1,\alpha}$ and $\mathcal{M}_{sig,\alpha}$ by setting

$$\mathcal{M}_{1,\alpha} = \{\mu \in \mathcal{M}_1 :\ m_\alpha(\mu) < \infty\},$$
$$\mathcal{M}_{sig,\alpha} = \{\mu \in \mathcal{M}_{sig} :\ m_\alpha(\mu) < \infty\}$$

where

$$m_\alpha(\mu) = \int\limits_X (\varrho(x,c))^\alpha |\mu|(dx).$$

Evidently $\mathcal{M}_{sig,\alpha} \subset \mathcal{M}_{sig,\beta}$ for $\alpha \geq \beta$. It is evident that these spaces do not depend on the choice of c. For every $f : X \to \mathbb{R}$ and $\mu \in \mathcal{M}_{sig}$ we write

$$\langle f,\mu\rangle = \int\limits_X f(x)\mu(dx), \tag{1}$$

whenever this integral exists. In the space $\mathcal{M}_1$ we introduce *the Kantorovich-Wasserstein metric* by the formula

$$\|\mu_1 - \mu_2\|_{\mathcal{K}} = \sup\{|\langle f, \mu_1 - \mu_2\rangle| : \ f \in \mathcal{K}\} \qquad \text{for} \qquad \mu_1, \mu_2 \in \mathcal{M}_1, \tag{2}$$

where $\mathcal{K}$ is the set of functions $f : X \to R$ which satisfy the condition

$$|f(x) - f(y)| \leq \varrho(x, y) \quad \text{for} \quad x, y \in X.$$

3. Markov operators

An operator $P : \mathcal{M} \to \mathcal{M}$ is called a *Markov operator* if it satisfies the following conditions:

(i) P is positively linear

$$P(\lambda_1\mu_1 + \lambda_2\mu_2) = \lambda_1 P\mu_1 + \lambda_2\mu_2$$

for $\lambda_1, \lambda_2 \geq 0$ and $\mu_1, \mu_2 \in \mathcal{M}$,

(ii) P preserves the measure of the space

$$P\mu(X) = \mu(X) \qquad \text{for} \qquad \mu \in \mathcal{M}. \tag{3}$$

Every Markov operator P can be uniquely extended as a linear operator to the space of signed measures. Namely for every $\mu \in \mathcal{M}_{sig}$ we define

$$P\mu = P\mu_1 - P\mu_2, \qquad \text{where} \qquad \mu = \mu_1 - \mu_2; \quad \mu_1, \mu_2 \in \mathcal{M}.$$

It is easy to verify that this definition of P does not depend on the choice of μ_1, μ_2.

A Markov operator P is called a *regular operator* if there exists an operator acting on the space of bounded Borel measurable functions $U : B(X) \to B(X)$ such that

$$\langle Uf, \mu\rangle = \langle f, P\mu\rangle \qquad \text{for} \quad f \in B(X),\ \mu \in \mathcal{M}. \tag{4}$$

The operator U is called *dual* to P.

If in addition $Uf \in C(X)$ for $f \in C(X)$, then the regular operator P is called the *Markov-Feller operator*.

Setting $\mu = \delta_x$ in (4) we obtain

$$(Uf)(x) = \langle f, P\delta_x\rangle \qquad \text{for} \quad f \in B(X), x \in X, \tag{5}$$

where $\delta_x \in \mathcal{M}_1$ is the point (Dirac) measure supported at x.

From formula (5) it follows immediately that U is linear and satisfies the following conditions

$$Uf \geq 0 \qquad \text{for} \qquad f \geq 0 \quad f \in B(X), \tag{6}$$

$$U\mathbf{1}_X = \mathbf{1}_X, \tag{7}$$

$$Uf_n \downarrow 0 \qquad \text{for} \qquad f_n \downarrow 0, \quad f_n \in B(X). \tag{8}$$

Here $f_n \downarrow 0$ means that the sequence (f_n) is decreasing and pointwise converges to 0.

Conditions (6)–(8) allow to reverse the role of P and U. Namely, if a linear operator U satisfying (6)–(8) is given we may define a Markov operator $P : \mathcal{M} \to \mathcal{M}$ by setting

$$P\mu(A) = \langle U\mathbf{1}_A, \mu\rangle \qquad \text{for} \qquad \mu \in \mathcal{M},\ A \in \mathcal{B}_X. \tag{9}$$

The dual operator U has a unique extension to the set of all Borel measurable nonnegative (not necessarily bounded) functions on X, such that formula (4) holds. Namely for a Borel measurable function $f : X \to \mathbb{R}^+$ we write

$$Uf(x) = \lim_{n\to\infty} Uf_n(x),$$

where $(f_n), f_n \in B(X)$ is an increasing sequence of bounded Borel measurable functions converging pointwise to f. Since the sequence (Uf_n) is increasing the limit Uf exists. Further from the Lebesgue monotone convergence theorem it follows that Uf satisfies (4). This formula shows that the limit is defined in a unique way and does not depend on the particular choice of the sequence (f_n). Evidently this extension is positively linear and monotonic.

4. Asymptotic stability of Markov operators.

A measure μ is called *stationary* (or *invariant*) with respect to a Markov operator P if $P\mu = \mu$.

A Markov operator is called *asymptotically stable* if there is a stationary distribution μ_* such that

$$\lim_{n\to\infty} \|P^n\mu - \mu_*\|_{\mathcal{K}} = 0 \qquad \text{for} \quad \mu \in \mathcal{M}_1. \tag{10}$$

A distribution μ_* satisfying (10) is unique.

For given $c \in X$ define

$$\varrho_c(x) := \varrho(x, c) \qquad \text{for} \qquad x \in X.$$

An important role in the study of the asymptotic behaviour of Markov-Feller operator P is played by the function $U\varrho_c$, where U denotes the dual operator to P. Since ϱ_c is continuous and nonnegative the function $U\varrho_c$ is well defined.

Using the Kantorovich-Rubinstein maximum principle, it is easy to give a sufficient condition for the asymptotic stability of Markov operators acting on the space of signed measures ([3]). As before let c be a fixed element of X and let $\varrho_c^\alpha(x) := \big(\varrho(x, c)\big)^\alpha$ for $x \in X$ and $\alpha > 0$.

Theorem 1. *Let $P : \mathcal{M}_{sig} \to \mathcal{M}_{sig}$ be a Markov–Feller operator and let U be its dual. Assume that*

$$|Uf(x) - Uf(y)| < \varrho(x, y) \qquad \textit{for} \quad x, y \in X,\ x \neq y, \tag{11}$$

and for every $f \in \mathcal{K}$.

Moreover, we assume that there exist constants $A, B \geq 0$, and $\alpha > 1$, such that

$$(U^n \varrho_c^\alpha)(x) \leq A\varrho_c^\alpha(x) + B \qquad \textit{for} \quad x \in X \ \textit{and}\ n = 0, 1, 2, \ldots. \tag{12}$$

Then P is asymptotically stable with respect to the Kantorovich-Wasserstein metric.

5. Discrete time stochastically perturbed dynamical system

We use this criterion to study stochastically perturbed dynamical systems. A discrete time dynamical system describes the evolution of points by means of one transformation $x_{n+1} = S(x_n)$. But practice this evolution can be randomly modified. In this case we have a family of transformations $\{S_y\}_{y\in Y}$ and at each step x_{n+1} is given by $S_y(x_n)$, where y is randomly selected. Such a process can be described by a stochastically perturbed dynamical system.

Let $(\Omega, \Sigma,\ \text{prob})$ be a probability space and let $(Y, \mathcal{A})$ be a measurable space. We consider a discrete time stochastically perturbed dynamical system on a locally compact separable space (X, ϱ) given by the recurrence formula:

$$x_{n+1} = S(x_n, \xi_n) \qquad \text{for} \quad n = 0, 1, \dots, \tag{13}$$

where $\xi_n : \Omega \to Y$ is a sequence of random elements, $S : X \times Y \to X$ is a given deterministic transformation and $S(x, y) := S_y(x)$. In our study of the asymptotic behaviour of (13) we assume that the following conditions are satisfied:

(i) The function S is measurable on the product space $X \times Y$ and for every fixed $y \in Y$ the function $S(\cdot, y)$ is continuous.
(ii) The random elements $\xi_0, \xi_1, \dots$ are independent and have the same distribution, i.e., the measure

$$\varphi(A) = \ \text{prob}(\xi_n \in A) \qquad \text{for} \quad A \in \mathcal{A},$$

is the same for all n.
(iii) The initial value $x_0 : \Omega \to X$ is a random element independent of the sequence (ξ_n).

It is easy to derive a recurrence formula for the measures

$$\mu_n(A) = \ \text{prob}(x_n \in A), \quad A \in \mathcal{B}(X),$$

corresponding to the dynamical system (13). Namely $\mu_{n+1} = P\mu_n$, $n = 0, 1, \dots$, where the operator $P : \mathcal{M}_1 \to \mathcal{M}_1$ is given by the formula

$$P\mu(A) = \int_X \Big(\int_Y 1_A(S(x, y))\varphi(dy)\Big)\mu(dx). \tag{14}$$

The operator P is a Markov–Feller operator and its dual U has the form

$$Uf(x) = \int_Y f(S(x, y))\varphi(dy) \qquad \text{for} \quad f \in C(X). \tag{15}$$

Now define a sequence of functions S_n by setting

$$S_1(x, y_1) = S(x, y_1), \qquad S_n(x, y_1, \dots, y_n) = S(S_{n-1}(x, y_1, \dots, y_{n-1}), y_n).$$

Using this notation we have

$$U^n f(x) = \int_Y \dots \int_Y f(S_n(x, y_1, \dots, y_n))\varphi(dy_1)\dots\varphi(dy_n).$$

Proposition 2. *Assume that the mapping $S : X\times Y \to X$ and the sequence of random elements (ξ_n) satisfy conditions* (i)–(iii). *Assume moreover that there is $n \in \mathbb{N}$ such that, the mathematical expectation of $\varrho(S(x,\xi_n), S(\overline{x},\xi_n)$ satisfies*

$$E(\varrho(S(x,\xi_n), S(\overline{x},\xi_n)) < \varrho(x,\overline{x}) \qquad \textit{for} \quad x,\overline{x} \in X, \quad x \neq \overline{x}, \tag{16}$$

and there exist constants $\alpha > 1$, A, $B \in \mathbb{R}^+$ such that

$$U^n \varrho_c^\alpha(x) \le A\varrho_c^\alpha(x) + B, \qquad \textit{for} \quad x \in X, \quad n = 0,1,2,..... \tag{17}$$

Then the operator P defined by (14) is asymptotically stable with respect to the Kantorovich-Wasserstein metric.

□

Using Proposition 2 it is easy to obtain a few known results concerning the stability of Markov operators.

In fact from Proposition 2 we immediately obtain as a special case the stability theorem of Lasota-Mackey (see [9], Theorem 2) where the conditions

$$E(|S(x,\xi_n) - S(z,\xi_n)|) < |x-z| \qquad \text{for} \quad x,z \in X \subset \mathbb{R}^d,\ x \neq z$$

and

$$E(|S(x,\xi_n)|^2) \le A|x|^2 + B \qquad \text{for} \quad x \in X \subset \mathbb{R}^d,$$

were assumed. The symbol $|\cdot|$ denotes an arbitrary, not necessary Euclidean, norm in $\mathbb{R}^d$ and A and B are nonnegative constants with $A < 1$.

Furthermore, in the case when X is a locally compact separable metric space, Proposition 2 contains a result of Łoskot and Rudnicki (see [10], Theorem 3). Namely, they proved the asymptotic stability of P if

$$\varrho(S(x,y), S(\overline{x},y)) \le \lambda(y)\varrho(x,\overline{x}) \qquad \text{for} \quad x,\overline{x} \in X$$

and

$$E\varrho_c(S(c,\xi_1)) < \infty,$$

where $\lambda : Y \to \mathbb{R}_+$ and $E\lambda(\xi_1) < 1$.

In the special case when, $Y = \{1,...,N\}$, the stochastic dynamical system (13) reduces to an *iterated function system*

$$(S_1,...,S_N; p_1,...,p_N) \quad \text{where} \quad S_k(x) = S(x,k) \quad \text{and} \quad p_k = \text{ prob}(\xi_n = k).$$

Now operators (14) and (15) have the form:

$$P\mu(A) = \sum_{k=1}^N p_k \mu(S_k^{-1}(A)) \qquad \text{and} \quad Uf(x) = \sum_{k=1}^N p_k f(S_k(x)). \tag{18}$$

We will assume the following conditions:

$$\sum_{k=1}^N p_k \varrho(S_k(x), S_k(\overline{x})) < \varrho(x,\overline{x}), \qquad \text{for} \quad x,\overline{x} \in X,\ x \neq \overline{x} \tag{19}$$

and

$$\varrho(S_k(x), c) \le L_k \varrho(x,c) \qquad \text{for} \quad x \in X,\ k = 1,\dots,N, \tag{20}$$

where c is a given point in X and the constants L_k are nonnegative constants.

In this case Proposition 2 implies the following result

Corollary 3. *If the IFS $(S_1,...,S_N;p_1,...,p_N)$ satisfies conditions (19), (20) and there exists a constant $\alpha>1$ such that*

$$\sum_{k=1}^{N} p_k L_k^{\alpha} < 1, \tag{21}$$

then this system is asymptotically stable.

6. Biological application: Dynamical systems with multiplicative perturbations

Our sequences (13) have no typical properties which are usually assumed in the known versions of the limit theorems. In particular, the random variables are not independent. The central limit theorem may be extended to various cases when the variables in the sum are not independent. We shall here only indicate one of these extensions for dynamical systems with multiplicative perturbations, which has a considerable importance for various applications, especially to biological problems. It will be convenient to use a terminology directly connected with some of the biological applications.

We now turn our attention to dynamical systems with multiplicative perturbations of the form

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \xi_{n+1} g(\mathbf{X}_n),$$

where $g:\mathbf{R}\to\mathbf{R}$ is integrable. Such dynamical systems have been considered in a biological context by H. Cramer in 20th century (1927) (see [2]) . It will be convenient to use a terminology directly connected with some of the biological applications. Suppose that we have n impulses $\xi_1,\ldots,\xi_n$ acting the order of their indices. These we consider as independent random variables. Denote by $\mathbf{X}_n$ the size of the organ which is produced by the impulses $\xi_1,\ldots,\xi_n$. We may then suppose e.g. that the increase caused by the impulse ξ_{n+1} is proportional to ξ_{n+1} and to some function $g(\mathbf{X}_n)$ of the momentary size the organ

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \xi_{n+1} * f\big(\mathbf{X}_n\big) \Longleftrightarrow \xi_{n+1} = \frac{\mathbf{X}_{n+1}-\mathbf{X}_n}{f\big(\mathbf{X}_n\big)}.$$

It follows that

$$\xi_1+\xi_2+\ldots+\xi_n = \sum_{k=0}^{n-1} \frac{\mathbf{X}_{k+1}-\mathbf{X}_k}{f\big(\mathbf{X}_k\big)}.$$

If each impulse only gives a slight contribution to the growth of the organ, we thus have approximately

$$\xi_1+\xi_2+\ldots+\xi_n \approx \int_{\mathbf{X}_0}^{\mathbf{X}} \frac{dt}{f\big(t\big)},$$

where $\mathbf{X}=\mathbf{X}_n$ denotes the final size of the organ. By hypothesis $\xi_1,\xi_2,\ldots,\xi_n$ are independent variables, and n may be considered as a large number. Under the general regularity conditions of the central random theorem it thus follows that, in the limit, the function of the random variable $\mathbf{X}$ appearing in the second member is normally distributed. Consider the case $f(t) = t$.

$$\int_{\mathbf{X}_0}^{\mathbf{X}} \frac{dt}{t} = \ln\, t\,\Big|_{\mathbf{X}_0}^{\mathbf{X}} = \ln\, \frac{\mathbf{X}}{\mathbf{X}_0}$$

The effect of each impulse is then directly proportional to the momentary size of the organ. In this case we thus fine that $\log \mathbf{X}$ is normally distributed.

If, more generally, $\log(\mathbf{X} - a)$ is normal (m, σ), it is easily seen that the variable $\mathbf{X}$ itself has the *logarithmico-normal distribution*:

$$\mathbf{f}_{\mathbf{X}}(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma\, x}\, e^{-\frac{(\ln\, x - m)^2}{2\sigma^2}} & \text{dla} \quad x > 0, \\ 0 & \text{dla} \quad x \leq 0. \end{cases}$$

References

[1] H. Cramer, *Metody matematyczne w statystyce*, PWN Warszawa 1958.

[2] H. Cramer, *Sannolikhetskalkylen och nagra av dess anvandingar* , Stockholm 1927.

[3] H. Gacki , *An application of the Kantorovich-Rubinstein maximum principle in the stability theory of Markov operators*, Bull. Pol. Ac. Math. 46 (1998), 215–223.

[4] H. Gacki, *Kantorovich-Rubinstein maximum principle in the stability theory of Markov semigroups*, Bull. Pol. Ac. Math. 52.2 (2004), 211-222.

[5] H. Gacki, *On the Kantorovich-Rubinstein maximum principle for the Fortet-Mourier norm*, Ann. Pol. Math. 86.2 (2005), 107-121.

[6] H. Gacki, *Applications of the Kantorovich-Rubinstein maximum principle in the theory of Markov semigroups*, Dissertationes Mathematicae 448 (2007), 1-59.

[7] H. Gacki and A. Lasota, *A nonlinear version of the Kantorovich-Rubinstein maximum principle*, Nonlinear Anal. 52 (2003), 117-125.

[8] , L. V. Kantorovich, G. S. Rubinstein, *On a space of completely additive functions* (in Russian), Vestnik Leningrad Univ. 13 **(1958)**, 52–59.

[9] A. Lasota and M.C. Mackey, *Stochastic perturbation of dynamical system: the weak convergence of measures*, J. Math. Anal. Appl. 138 (1989), 232-248.

[10] K. oskot and R. Rudnicki, *Limit theorems for stochastically perturbated dynamical systems*, J. Appl. Prob. 32 (1995), 459-469.

[11] G. Monge, *Mémoire sur la théorie des déblais et des ermblais*, Histoire de l'Académie des sciences de Paris, avec les Mémoires de mathématique et de phisique pour la même année, pp. 666–704, avec 2 pl. **(1781)**.

[12] S. T. Rachev and R. M. Shortt, *Duality theorems for Kantorovich-Rubinstein and Wasserstein functionals*, Dissertationes Mathematicae CCXCIX (1990), 1-35.

[13] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models* , John Wiley and Sons, New York 1991.

[14] C. Villani, *Optimal transport, old and new*, Springer-Verlag, Berlin 2006.

Henryk Gacki, Institute of Mathematics, Silesian University, Bankowa 14, 40-007 Katowice, Poland)

E-mail address: gacki@ux2.math.us.edu.pl