On Kalman models over a commutative ring

Vakhtang Lomadze
Institute of Mathematics
Tbilisi 380093, Georgia

Abstract

There is a good notion of rational functions with coefficients in a commutative ring. Using this notion, we easily obtain a neat generalization of Chapter 10 of the classical book by Kalman *et al.* to linear systems over an arbitrary commutative ring. The generalizations certainly exist already. However, we believe that the approach we present is more natural and straightforward.

In this note we would like to show that the classical results of Kalman given in [Ch. 10, 5] can be generalized with absolutely no difficulty to the case when linear systems are defined over a commutative ring. For other generalizations we refer to [2, 3, 4, 6, 7].

Throughout, A is a commutative ring (with a unit of course, but not necessarily Noetherian), s an indeterminate, m an input number and p an output number.

By a monic polynomial we shall understand a one whose leading coefficient is an invetible element of A. Obviously monic polynomials form a multiplicative subset in $A[s]$. The corresponding localization of $A[s]$, denoted by $A(s)$, is called the ring of rational functions with coefficients in A. Thus, by definition, $A(s)$ consists of fractions of the form f/g, where f is an arbitrary polynomial and g a monic polynomial. We remark that a monic polynomial can not be a zero divisor. Hence, $f_1/g_1 = f_2/g_2$ if and only if $f_1 g_2 = f_2 g_1$. This implies, in particular, that the canonical homomorphism

$$A[s] \to A(s), \quad f \mapsto f/1$$

is an embedding, and we shall identify $A[s]$ with its image under this embedding.

A rational function f/g is called proper if $deg(f) \leq deg(g)$. It is easily seen that proper rational functions form a ring, and we shall denote it by O. Clearly, we have $A(s) = \cap_{n \geq 0} s^n O$. Notice that the Euclidean algorithm holds; namely, if f and g are polynomials and if g is monic, then there exists a unique pair of polynomials q and r such that $f = qg + r$ and $deg(r) < deg(g)$. (We put $deg(0) = -\infty$.) Consequently, we have a fundamental relation

$$A(s) = A[s] \oplus s^{-1} O.$$

Given a rational function f, define its residue $Res(f)$ as the coefficient at s^{-1} in the representation of f as a series in $A((s^{-1}))$. The map $Res : A(s) \to A$ is A-linear and vanishes on $A[s]$. Therefore it determines a canonical A-linear map $A(s)/A[s] \to A$, which will be denoted by Res again.

We make the following

Convention. An element of an $A[s]$-module will be said to be torsion if (and only if) it is anihilated by a monic polynomial.

Finitely generated torsion modules over $A[s]$ will be called finite.

Let X be a finitely generated module over A, and let F be an endomorphism of X. Clearly, X together with the multiplication

$$A[s] \times X \to X, \quad (a, x) \mapsto a(F)(x)$$

becomes a module over $A[s]$. Let X_F denote this module.

Lemma 1. *If X and F are as above, then X_F is a finite $A[s]$-module.*

Proof. Obviously X_F is finitely generated. Consider any epimorphism $X' \to X$, where X' is a free module of finite rank. Clearly F can be "lifted" to an endomorphism F' of X'. If $h(s)$ is the characteristic polynomial of F', then, by the Cayley-Hamilton theorem, $h(F') = 0$. It follows that $h(F) = 0$. So, $h(s)x = 0$ for each $x \in X$. $\square$

Let X and F be as above, and let $X[s] \to X_F$ be the $A[s]$-homomorphism taking $\sum(x_i \otimes s^i)$ to $\sum F^i(x_i)$.

Lemma 2. *The sequence*

$$0 \to X[s] \overset{sI-F}{\to} X[s] \to X_F \to 0$$

is exact.

Proof. The statement can be found in [1] (see Prop. 18, §8, Ch III), except that $X[s] \to X[s]$ is injective. So we restrict ourselves by proving this.

Suppose that $(x_0 \otimes 1) + (x_1 \otimes s) + \cdots + (x_n \otimes s^n)$ goes to zero, that is,

$$-Fx_0 \otimes 1 + (x_0 - Fx_1) \otimes s + \cdots + (x_{n-1} - Fx_n) \otimes s^n) + x_n \otimes s^{n+1} = 0.$$

It immediately follows from this that $x_n = 0,\ x_{n-1} = 0,\ \ldots, x_0 = 0.$ $\square$

Corollary 1. *The homomorphism*

$$sI - F : X(s) \to X(s)$$

is bijective.

Proof. Tensoring the exact sequence of the previous lemma by $A(s)$, we get an exact sequence

$$0 \to X(s) \overset{sI-F}{\to} X(s) \to X_F \otimes A(s) \to 0.$$

Because X_F is a torsion module, we have $X_F \otimes A(s) = 0$. The statement follows. $\square$

If X is an A-module, then clearly $X[s]$ is a torsion free $A[s]$-module. Hence, we may identify $X[s]$ with its image under the canonical homomorphism $X[s] \to X(s)$.

If X and Y are finitely generated A-modules and $R : X[s] \to Y[s]$ is an $A[s]$-homomorphism, then we denote by R again the induced $A(s)$-homomorphism $X(s) \to Y(s)$. We define

$$RX[s] = \{Rx|\ x \in X[s]\} \quad \text{and} \quad R^{-1}Y[s] = \{x \in X(s)|\ Rx \in Y[s]\}.$$

Let X, X_1 and X_2 be finitely generated A-modules. We say that $A[s]$-homomorphisms $G_1 : X_1[s] \to X[s]$ and $G_2 : X_2[s] \to X[s]$ are left coprime if

$$G_1X_1[s] + G_2X_2[s] = X[s].$$

The definition of right coprimeness is less obvious. We say that $A[s]$-homomorphisms $H_1 : X[s] \to X_1[s]$ and $H_2 : X[s] \to X_2[s]$ are right coprime if

$$H_1^{-1}X_1[s] \cap H_2^{-1}X_2[s] = X[s].$$

Lemma 3. *The mapping $(X, F) \mapsto X_F$ establishes a one-to-one correspondence between the class of pairs consisting of a finitely generated A-module and its endomorphism and the class of finite $A[s]$-modules.*

Proof. Let Q be a finite $A[s]$-module. We can find a monic polynomial h that anihilates the whole Q. Clearly Q can be viewed as a module over $A[s]/hA[s]$. By the Euclidean algorithm, the elements $1, s, \dots , s^{d-1}$, where d is the degree of h, generate $A[s]/hA[s]$ as an A-module. It follows that Q is finitely generated as an A-module. One easily completes the proof. $\square$

Let X be a finitely generated A-module and $F : X \to X$ its endomorphism. For each linear map $G : A^m \to X$, define an $A[s]$-homomorphism

$$G_F : A[s]^m \to X_F, \quad G_F(\sum_{i\geq 0} u_i s^i) = \sum_{i\geq 0} F^i G(u_i).$$

Lemma 4. *The mapping $G \mapsto G_F$ establishes a one-to-one correspondence between*

$$Hom_A(A^m, X) \quad \textit{and} \quad Hom_{A[s]}(A[s]^m, X_F).$$

Proof. Obvious. $\square$

Let again X be a finitely generated A-module and $F : X \to X$ its endomorphism. For each linear map $H : X \to A^p$, define an $A[s]$-homomorphism

$$H_F : X_F \to A(s)^p/A[s]^p, \quad H_F(x) = (\sum_{i\geq 0} HF^i(x)s^{-i-1}) \text{mod} A[s]^p.$$

(This is well-defined because $\sum_{i\geq 0} HF^i(x)s^{-i-1} = H(sI - F)^{-1}x$.)

Lemma 5. *The mapping $H \mapsto H_F$ establishes a one-to-one correspondence between*

$$Hom_A(X, A^p) \quad \textit{and} \quad Hom_{A[s]}(X_F, A(s)^p/A[s]^p).$$

Proof. If $H : X \to A^p$ is a linear map, then obviously $H = Res \circ H_F$. Hence, the mapping is injective.

Let now ψ be a homomorphism of X_F into $A(s)^p/A[s]^p$, and let $H = Res \circ \psi$. Take an arbitrary $x \in X$, and suppose that

$$\psi(x) = (\sum_{i\geq 0} a_i s^{-i-1}) \mathrm{mod} A[s]^p.$$

Then

$$\forall n \geq 0, \quad \psi(s^n x) = (a_n s^{-1} + a_{n+1} s^{-2} + \cdots) \mathrm{mod} A[s]^p.$$

We see that

$$\forall n \geq 0, \quad a_n = Res\psi(s^n x) = HF^n x.$$

Hence, $\psi = H_F$. $\square$

We can now pass to system theory.

A transfer function is defined to be a rational matrix of size $p \times m$. A linear system is a quintuple $(X; F, G, H, J)$, where X is a finitely generated A-module $F : X \to X$, $G : A^m \to X$, $H : X \to A^p$ are A-linear maps, and J is a polynomial $p \times m$ matrix. If J is constant, then the system is called regular. The rational matrix $H(sI - F)^{-1}G + J$ is called the transfer function.

Let $(X; F, G, H, J)$ and $(X_1; F_1, G_1, H_1, J_1)$ be linear systems. If $J = J_1$, then a transformation of the first one into the other is a linear map $K : X \to X_1$ such that $F_1 K = KF$, $G_1 = KG$ and $H = H_1 K$. A transformation is not defined when $J \neq J_1$.

Let (X, F, G, H, J) be a linear system. We remind that

$$X[s]/(sI - F)X[s] \simeq X_F \tag{1}$$

(see Lemma 2), and that

$$F^n = a_1 F^{n-1} + \cdots + a_n I \tag{2}$$

for some $n \in \mathbb{Z}_+$ and $a_1, \ldots, a_n \in A$ (see the proof of Lemma 1).

Theorem 1. *The following conditions are equivalent:*

(a) $Im(G) + Im(FG) + \cdots + Im(F^n G) = X$ for sufficiently large n;

(b) $G_F : A[s]^m \to X_F$ is surjective;

(c) $G : A[s]^m \to X[s]$ and $sI - F : X[s] \to X[s]$ are left coprime.

Proof. $(a) \iff (b)$ Obviously $ImG_F = \sum_{i\geq 0} ImF^i G$. From this and (2) the equivalence follows.

$(b) \iff (c)$ Using (1), we have a commutative square

$$\begin{array}{ccc} A[s]^m & = & A[s]^m \\ \downarrow & & \downarrow \\ X[s]/(sI-F)X[s] & \simeq & X_F \end{array},$$

where the left downward arrow is defined as $u \mapsto (Gu) \mathrm{mod}(sI - F)X[s],\ u \in A[s]^m$. Clearly, (c) means surjectivity of this arrow, and the equivalence follows. $\square$

Theorem 2. *The following conditions are equivalent:*

(a) $Ker(H) \cap Ker(HF) \cap \cdots \cap Ker(HF^n) = 0$ *for sufficiently large* n*;*

(b) $H_F : X_F \to A(s)^p/A[s]^p$ *is injective;*

(c) $H : X[s] \to A[s]^p$ *and* $sI - F : X[s] \to X[s]$ *are right coprime.*

Proof. $(a) \iff (b)$ Obviously $KerH_F = \cap_{i\geq 0} KerHF^i$. From this and (2) the equivalence follows.

$(b) \iff (c)$ Using (1), we have a commutative square

$$\begin{array}{ccc} X[s]/(sI-F)X[s] & \simeq & X_F \\ \downarrow & & \downarrow \\ A(s)^p/A[s]^p & = & A(s)^p/A[s]^p \end{array},$$

where the left downward arrow is induced by $x \mapsto H(sI-F)^{-1}x,\ x \in X[s]$. Clearly, (c) means injectivity of this arrow, and the equivalence follows. $\square$

A linear system $(X; F, G, H, J)$ is said to be controllable if it satisfies the conditions of Theorem 1 and observable if it satisfies the conditions of Theorem 2. A linear system is called minimal if it is both controllable and observable.

A Kalman model is a quadruple $(Q; \phi, \psi, J)$, where Q is a finite $A[s]$-module, $\phi : A[s]^m \to Q$, $\psi : Q \to A(s)^p/A[s]^p$ are $A[s]$-homomorphisms and $J : A^m \to A^p$ is a polynomial matrix of size $p \times m$. If J is constant, the model is called regular. The model is controllable if ϕ is surjective, and observable if ψ is injective. A Kalman model is called minimal if it is both controllable and observable.

Let $(Q; \phi, \psi, J)$ and $(Q_1; \phi_1, \psi_1, J_1)$ be Kalman models. If $J = J_1$, then a transformation of the first one into the other is a homomorphism $\theta : Q \to Q_1$ such that $\phi_1 = \theta\phi$ and $\psi = \psi_1\theta$. A transformation is not defined when $J \neq J_1$.

Given a linear system $\Sigma = (X; F, G, H, J)$, we set

$$\mathrm{K}(\Sigma) = (X_F; G_F, H_F, J).$$

This is a Kalman model.

It is easily seen that "K" is a functor. Obviously, this functor preserves the property of regularity. The theorems 1 and 2 say that this functor preserves also the properties of controllability and observability.

Theorem 3. *The functor* "K" *establishes a canonical equivalence between linear systems and Kalman models.*

Proof. Follows from the lemmas 3, 4 and 5. $\square$

Corollary 2. *There is a canonical one-to-one correspondence between isomorphism classes of minimal linear systems and transfer functions.*

Proof. Given a transfer function T, let $Q(T)$ denote the image under the composition

$$A[s]^m \xrightarrow{T} A(s)^p \to A(s)^p/A[s]^p,$$

and let T_{spr} and T_{pol} denote respectively the strictly proper and the polynomial parts of T. Then, clearly

$$(Q(T); T_{spr}, id, T_{pol})$$

is a minimal Kalman model. It is easily seen that (up to isomorphism) every minimal Kalman model is obtained this way. The statement follows now from the previous theorem. □

References

[1] N. Bourbaki, *Elements of Mathematics, Algebra I*, Springer-Verlag, Berlin Heidelberg New York, 1989.

[2] J.W. Brewer, J.W. Bunce, and F.S. Van Vleck, *Linear Systems over Commutative Rings*, Marsel Dekker, New York, 1986.

[3] S. Eilenberg, *Automata, Languages, and Machines*, Acad. Press., New York, 1974.

[4] G. Conte and A.M. Perdon, Systems over a principal ideal domain. A polynomial model approach, SIAM J. Contr. Optim., 20 (1982), 112-124.

[5] R.E. Kalman, P.L. Falb, and M.A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

[6] E.D. Sontag, *Linear systems over commutative rings: a survey*, Ricerche di Automatica 7 (1976), 1-34.

[7] B. Wyman, *Linear Systems over Commutative Rings*, Stanford University, Stanford, CA, 1972.