Synthesis of generating filter with specified mean anisotropy level

Kustov A.Yu.

ICS RAS

June, 17-23, 2012

- *Владимиров И.Г., Даймонд Ф., Клоеден П.* Анизотропийный анализ робастного качества линейных нестационарных дискретных систем на конечном интервале времени// АиТ, 2006.
- *Cover T.M., Thomas J.A.* Elements of Information Theory, 1991.
- *Diamond P., Vladimirov I., Kurdjukov A., Semyonov A.* Anisotropy-based Performance Analysis of Linear Discrete Time Invariant Control Systems.// *International journal of control, Vol.74, No.1*, 2001.
- *Gray R.M.* Entropy and Information Theory. *Springer-Verlag, New York*, 1990.

Relative entropy

Relative entropy (or Kulback-Leibler divergence) is a measure of the difference between two stochastic variables W, V with corresponding probability density functions f and g, and is defined by

$$\mathbf{D}(f||g) = \mathbf{E}_f\left[\ln\left(\frac{f}{g}\right)\right] = \int_{\mathbb{R}^m} f(x)\ln\frac{f(x)}{g(x)}d\mathcal{V}. \tag{1}$$

In case of

$$g(x) = p_{m,\lambda}(x) = \frac{1}{\sqrt{(2\pi)^m\lambda}}\mathrm{e}^{-\frac{1}{2\lambda}x^{\mathrm{T}}x} \tag{2}$$

one have

$$\mathbf{D}(f||p_{m,\lambda}) = -h(W) + \frac{m}{2}\ln(2\pi\lambda) + \frac{\mathbf{E}[|W|^2]}{2\lambda},$$

where

$$h(W) = -\int_{\mathbb{R}^m} f(x)\ln f(x)d\mathcal{V}$$

is a differential entropy of W, and

$$\mathbf{E}[|W|^2] = \mathrm{tr}(\mathrm{cov}(W))$$

is a trace of its covariance matrix.

Anisotropy

Anisotropy of stochastic signal W is defined by

$$\mathbf{A}(W) = \min_{\lambda>0} \mathbf{D}(f||p_{m,\lambda}). \quad (3)$$

Anisotropy of W describes measure of the difference between W and class of signals with normal distribution (2), i.e. between W and class of gaussian "white" noises.

In one-dimensional case with p.d.f.

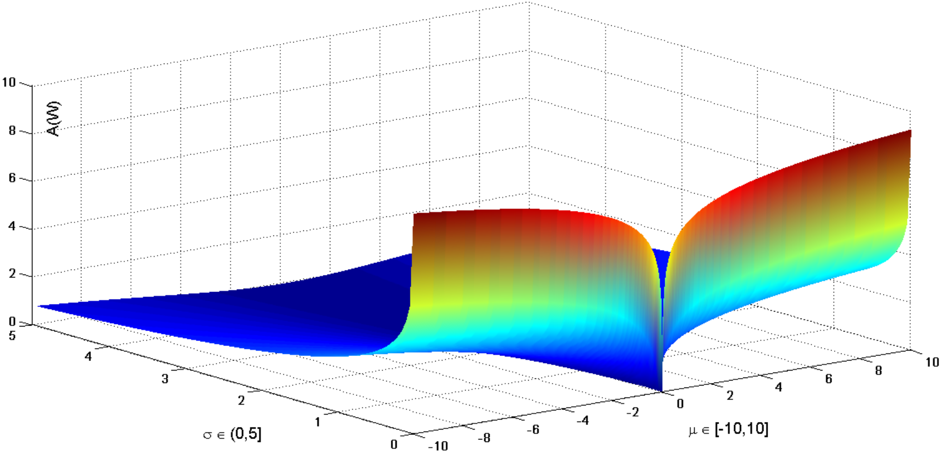
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

of signal W its anisotropy connected with parameters μ and σ by the relationship

$$\mathbf{A}(W) = \ln \sqrt{1 + \left(\frac{\mu}{\sigma}\right)^2}.$$

Illustration of this formula is given on the next slide.

Anisotropy (illustration)



Mean anisotropy

Let $\{w_k\}$ be stationary sequence of stochastic m-dimensional square-integrable vectors, i.e.

$$w_k \in \mathbb{R}^m, \quad \mathbf{E}[|w_k|^2] < +\infty.$$

Mean anisotropy of such a sequence is defined by

$$\overline{\mathbf{A}}(W) = \lim_{N \to +\infty} \frac{\mathbf{A}(W_{0:N-1})}{N}, \tag{4}$$

where

$$W_{0:N-1} = \begin{bmatrix} w_0 \\ \vdots \\ w_{N-1} \end{bmatrix}.$$

Let sequence elements w_k are formed by the linear combination of previous gaussian "white" noises (i.e. by the instrumentality of filter G):

$$w_j = \sum_{k=0}^{\infty} g_k v_{j-k}, \quad j \in \mathbb{Z},$$

or by the system

$$\begin{cases} x_{k+1} = Ax_k + Bv_k, \\ w_k = Cx_k + Dv_k, \end{cases} \tag{5}$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{m \times m}$.

Filter interpretation

$$w_j = \sum_{k=0}^{\infty} g_k v_{j-k}, \quad j \in \mathbb{Z},$$

$$\begin{bmatrix} x_{k+1} \\ w_k \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_k \\ v_k \end{bmatrix},$$

$$W(z) = G(z)V(z).$$

Calculation of mean anisotropy

Theorem (Diamond et al. 2001)

For a given generating filter G with state-space performance (5) ($\rho(A) < 1$, $\det D \neq 0$) mean anisotropy has the appearance

$$\overline{\mathbf{A}}(W) = -\frac{1}{2} \ln \det \left(\frac{m\Theta}{\mathrm{tr}(CPC^T + DD^T)} \right), \tag{6}$$

where Θ is defined by the Riccati equation solution $R > 0$:

$$\begin{aligned} R &= ARA^T + BB^T - \Lambda\Theta\Lambda^T, & (7) \\ \Lambda &\doteq (ARC^T + BD^T)\Theta^{-1}, & (8) \\ \Theta &\doteq CRC^T + DD^T, & (9) \end{aligned}$$

and $P > 0$ is a controllability gramian satisfying the Lyapunov equation

$$P = APA^T + BB^T. \tag{10}$$

Problem statement

Let $F \in \mathcal{H}_\infty^{p\times m}$ be stable linear system. It is demand to find a solution of problem (6)–(10) in radicals, i.e. find an explicit dependence between $\overline{\mathbf{A}}(W)$ and matrices A, B, C, D allowing to compute partial derivatives

$$\frac{\partial \overline{\mathbf{A}}(W)}{\partial x}, \quad x \in \{a_{ij}, b_{ik}, c_{li}, d_{lm}\}.$$

It is easy to notice that mean anisotropy in (6) depends on determinant of certain matrix. It brings us to necessity of computation of determinant derivatives. If $A(t)$ is $n \times m$-matrix then

$$\frac{d \det A(t)}{dt} = \sum_{i=1}^{n} \det A_{io}(t) = \sum_{j=1}^{m} \det A_{oj}(t),$$

where $A_{io}(t)$ is matrix $A(t)$ with differentiated i-th row, and $A_{oj}(t)$ is matrix $A(t)$ with differentiated j-th column.

Such procedure doesn't simplify anything, therefore...

Separation of variables

Let A, B, C, D be block-diagonal matrices, i.e.

$$A = \mathrm{diag}\{A_1, \ldots, A_r\}, \quad B = \mathrm{diag}\{B_1, \ldots, B_r\},$$
$$C = \mathrm{diag}\{C_1, \ldots, C_r\}, \quad D = \mathrm{diag}\{D_1, \ldots, D_r\}.$$

Thus a pack of systems is given:

$$\begin{cases} x_{s,k+1} = A_s x_{s,k} + B_s v_{s,k}, \\ w_{s,k} = C_s x_{s,k} + D_s v_{s,k}, \end{cases}$$

where each one of them describes just a part of total output w_k. In this case equation (6) takes the next form:

$$\overline{\mathbf{A}}(W) = -\frac{1}{2} \ln \frac{\prod_{s=1}^{r} \det\left(\Sigma_s + \Xi_s\right)}{\left(\frac{1}{m} \sum_{s=1}^{r} \mathrm{tr}(\Sigma_s)\right)^m}, \tag{11}$$

where matrices Σ_s and Ξ_s are connected with the Riccati and Lyapunov equations by some rule.

Corollary

Under the assumption that some matrices A_s, B_s, C_s, D_s have diagonal structure equation (6) subject to (7)–(10) takes view

$$\overline{\mathbf{A}}(W) = -\frac{1}{2} \ln \frac{\prod_{s=1}^{r_1-1} (\sigma_s + \xi_s) \prod_{s=r_1}^{r} \det(\Sigma_s + \Xi_s)}{\left(\frac{1}{m} \left(\sum_{s=1}^{r_1-1} \sigma_s + \sum_{s=r_1}^{r} \operatorname{tr}(\Sigma_s) \right) \right)^m}, \tag{12}$$

where

$$\sigma_s = t_s^2 + \frac{p_s^2}{1 - a_s^2}, \quad \xi_s = \frac{-(\varkappa_s^2 + p_s^2) + |\varkappa_s^2 - p_s^2|}{2(1 - a_s^2)}, \quad \varkappa_s = (1 - a_s^2)t_s + a_s p_s.$$

Finally, lets consider the case of some unvarying matrices.

Final formula

Let matrices Σ_s and Ξ_s are unvarying with $s = \overline{r_1, r}$. Then formula (12) takes the following form:

$$\overline{\mathbf{A}}(W) = -\frac{1}{2}\ln c - \frac{1}{2}\ln \frac{\prod_{s=1}^{r_1-1}(\widetilde{\sigma}_s + \widetilde{\xi}_s)}{\left(\sum_{s=1}^{r_1-1}\widetilde{\sigma}_s + 1\right)^m}, \tag{13}$$

where

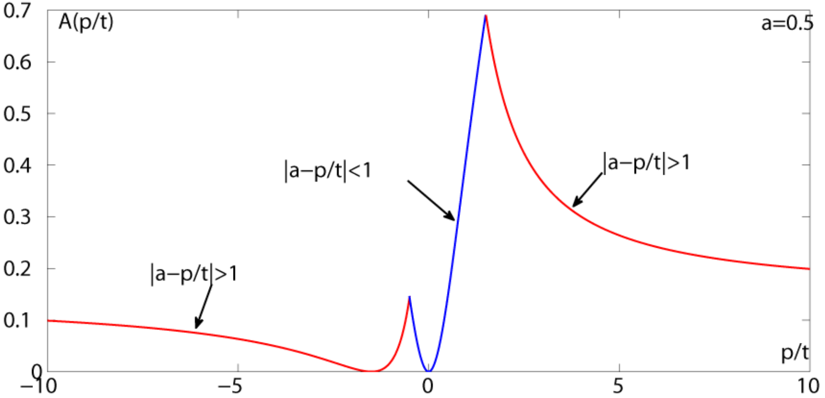
$$c = \frac{m^m}{\widetilde{c}^{m-r_1+1}} \prod_{s=r_1}^{r} \det(\Sigma_s + \Xi_s)$$

and

$$\widetilde{\sigma}_s = \sigma_s/\widetilde{c}, \quad \widetilde{\xi}_s = \xi_s/\widetilde{c}, \quad \widetilde{c} = \sum_{s=r_1}^{r} \mathrm{tr}(\Sigma_s),$$

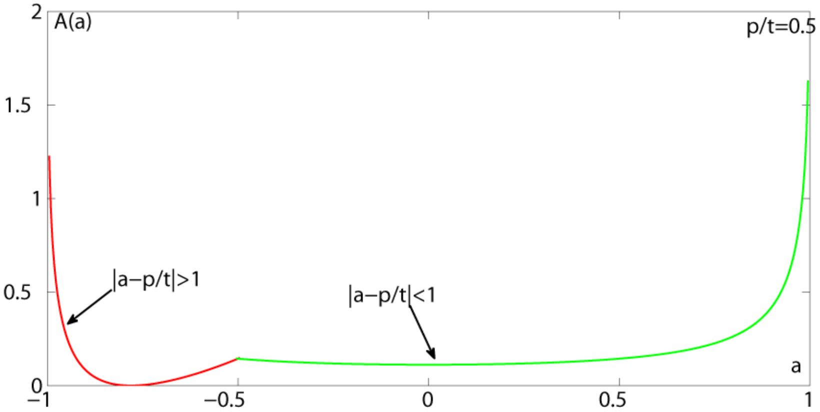
$$\sigma_s = t_s^2 + \frac{p_s^2}{1-a_s^2}, \quad \xi_s = \frac{-(\varkappa_s^2 + p_s^2) + |\varkappa_s^2 - p_s^2|}{2(1-a_s^2)}, \quad \varkappa_s = (1-a_s^2)t_s + a_s p_s.$$

One-dimensional case (a)



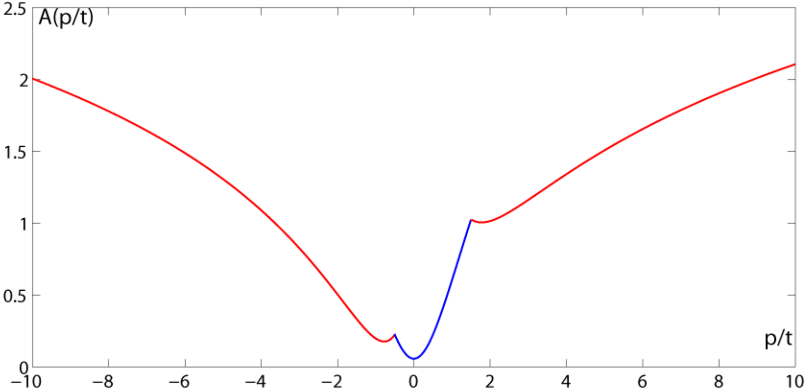
Pic.1 Graph of function $\overline{\mathbf{A}}_W(p/t)$

One-dimensional case (b)



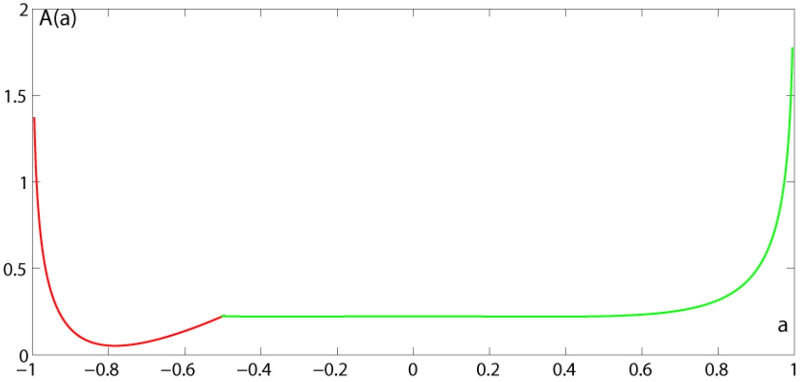
Pic.2 Graph of function $\overline{\mathbf{A}}_W(a)$

Two-dimensional case (a)



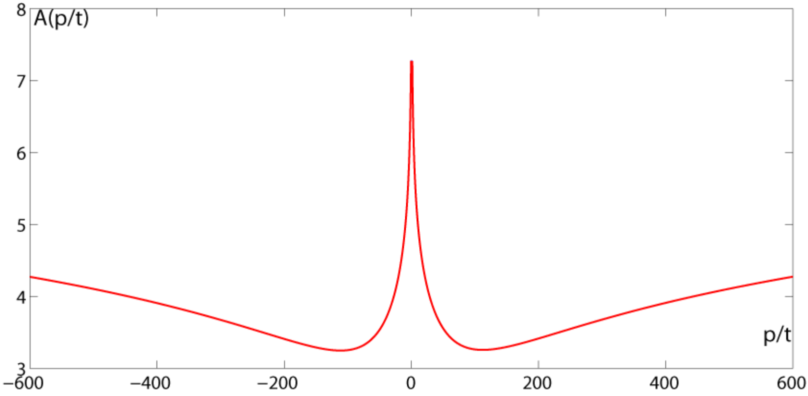
Pic.3 Graph of function $\overline{\mathbf{A}}_W(p/t)$

Two-dimensional case (b)



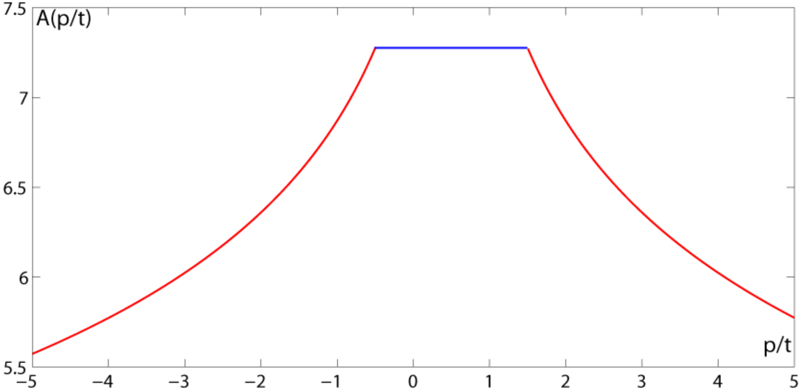
Pic.4 Graph of function $\overline{\mathbf{A}}_W(a)$

Two-dimensional case (b)



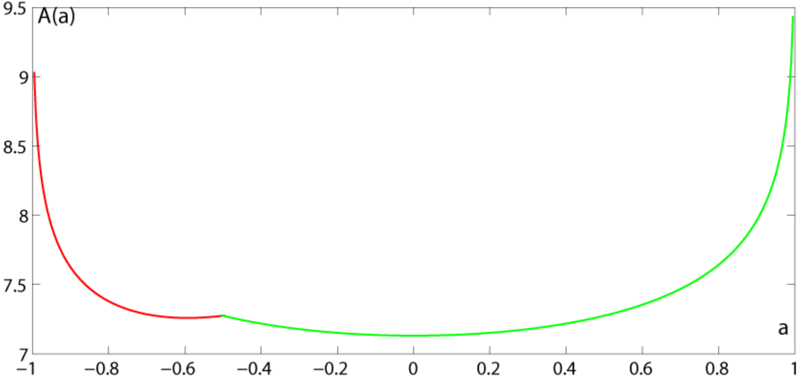
Pic.4 Graph of function $\overline{\mathbf{A}}_W(p/t)$

Two-dimensional case (b)



Pic.4 Graph of function $\overline{\mathbf{A}}_W(p/t)$

Two-dimensional case (b)



Pic.4 Graph of function $\overline{\mathbf{A}}_W(a)$

Application

Consider a LTI plant with state-space realization

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k, \\ y_k &= Cx_k + Dw_k, \end{aligned}$$

where x, y, w denote state, measured output and noise input respectively, and matrices A, B, C, D are from COMPleib test model 'AC4' (autopilot for an air-to-air missile):

$$A = \begin{bmatrix} 0.9996 & 0.0005 & -0.0001 & 0 \\ 0.0045 & 1 & -0.063 & 0 \\ 0 & 0 & 0.9277 & 0 \\ -0.0005 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.0005 \end{bmatrix},$$

$$C = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 1 \\ 0.01 & 0 \end{bmatrix}.$$

It is obvious that 2nd component of noise input does not effect on the subject state. Hence there is an ability to generate inputs $w_{k,1}$ and $w_{k,2}$ independently and as a result obtain an implicit dependence between mean anisotropy and coefficients of matrices corresponds to $w_{k,2}$.

Conclusion

Further researches:

- Problem of finding an explicit dependence between controller and generating filter matrices;
- Problem of mean anisotropy identification via observations;
- Problem of controller design under assumption of changeable environment, i.e. with varying mean anisotropy.

Conclusion:

- An explicit dependence between mean anisotropy and filter matrices is obtained; it is shown that obtaining of specified mean anisotropy level is connected with varying certain matrix coefficients.

THANKS FOR ATTENTION!