Applications of Spaces with Indefinite Metrics in Estimation (and Control)

Babak Hassibi

Department of Electrical Engineering
California Institute of Technology

Mathematical Theory of Networks and Systems
August 12-16

University of Notre Dame, South Bend, IN

Outline

- **H^2 and H^∞ Estimation**
- **Linear Estimation Krein Spaces**
  - projections in Krein space
  - Krein space Kalman filter
- **Applications**
  - risk-sensitive, game-theoretic, etc. types of estimation
  - square-root array algorithms
  - convergence of Riccati recursions
  - KYP lemma
- **Conclusion**

Least-Mean-Squares (or H^2) Estimation

Consider the (possibly) time-variant state-space model

$$\begin{cases} \mathbf{x}_{i+1} &= F_i \mathbf{x}_i + G_i \mathbf{u}_i \\ \mathbf{y}_i &= H_i \mathbf{x}_i + \mathbf{v}_i \end{cases} \qquad i \geq 0$$

where the disturbances $\{\mathbf{u}_i\}$ and $\{\mathbf{v}_i\}$ and the initial state $\mathbf{x}_0$ are assumed to be zero-mean *stochastic* processes such that

$$E \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \\ \mathbf{x}_0 \end{bmatrix}^* \begin{bmatrix} \mathbf{u}_j \\ \mathbf{v}_j \\ \mathbf{x}_0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} Q_i & 0 \\ 0 & R_i \end{bmatrix} \delta_{ij} & 0 \\ 0 & \Pi_0 \end{bmatrix}.$$

Problem: Estimate the signal $\mathbf{s}_i = L_i \mathbf{x}_i$, given the current and past observations $\mathbf{y}_i, \ldots, \mathbf{y}_0\}$ under the criterion

$$\min_{\hat{\mathbf{s}}_{i|i}} E \left(\mathbf{s}_i - \hat{\mathbf{s}}_{i|i} \right)^* \left(\mathbf{s}_i - \hat{\mathbf{s}}_{i|i} \right).$$

Solution: The celebrated Kalman filter (1960)

$$\mathbf{s}_{i|i} = L_i \hat{\mathbf{x}}_{i|i},$$

where

$$\hat{\mathbf{x}}_{i+1|i+1} = F_i \hat{\mathbf{x}}_{i|i} + K_{f,i+1}(\mathbf{y}_{i+1} - H_{i+1} F_i \hat{\mathbf{x}}_{i|i}), \quad \hat{\mathbf{x}}_0$$

with

$$K_{f,i} = P_i H_i^* R_{e,i}^{-1} \quad , \quad R_{e,i} = R_i + H_i P_i H_i^*$$

and where P_i satisfies the *Riccati recursion*

$$P_{i+1} = F_i P_i F_i^* + G_i Q_i G_i^* - F_i P_i H_i^* R_{e,i}^{-1} H_i P_i F_i^*, \quad P_0 = \Pi_0.$$

- This solution has been intensively studied and widely applied over the last 40 years.

H^∞ Estimation

Because of modeling errors and statistical uncertainties, it has been proposed to model the disturbances as *unknown, but non-random*:

$$\left\{\begin{array}{rcl} x_{i+1} & = & F_i x_i + G_i u_i \\ y_i & = & H_i x_i + v_i \\ s_i & = & L_i x_i \end{array}\right. \qquad i \geq 0$$

To guarantee robustness we now seek an estimator that bounds the maximum energy gain from the disturbances to the estimation errors by a given tolerance, γ^2, i.e.,

$$\sup_{\{u_i\},\{v_i\}\in h^2, x_0} \frac{\|s-\hat{s}\|_2^2}{x_0^* \Pi_0^{-1} x_0 + \|u\|_2^2 + \|v\|_2^2} < \gamma^2.$$

Problem: Determine whether it is possible to find such a level-γ estimator, and if so, find one.

Solution: (Khargonekar & Nagpal, Shaked & Yaesh, Green & Limebeer)

$$\hat{s}_{i|i} = L_i \hat{x}_{i|i},$$

where

$$\hat{x}_{i+1|i+1} = F_i \hat{x}_{i|i} + K_{s,i+1}(y_{i+1} - H_{i+1} F_i \hat{x}_{i|i}), \quad \hat{x}_0$$

with

$$K_{s,i} = P_i H_i^* (I + H_i P_i H_i^*)^{-1},$$

and where P_i satisfies the Riccati recursion

$$P_{i+1} = F_i P_i F_i^* + G_i G_i^* - F_i P_i \begin{bmatrix} H_i^* & L_i^* \end{bmatrix} R_{e,i}^{-1} \begin{bmatrix} H_i \\ L_i \end{bmatrix} P_i F_i^*, \quad P_0 = \Pi_0$$

with $R_{e,i} = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} H_i \\ L_i \end{bmatrix} P_i \begin{bmatrix} H_i^* & L_i^* \end{bmatrix}$. Moreover, the above are true, i.e., a level-γ H^∞ estimator exists, iff, the matrices

$$R_i = I \oplus (-I) \quad \text{and} \quad R_{e,i}$$

have the same inertia for all $i \geq 0$.

Remarks

This looks very much like the Kalman filter solution, except that

- we gave indefinite *covariance* matrices, $\begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix}$.
- the L_i (of the quantity to be estimated) enters the Riccati recursion, so that the structure of the filter depends on the linear combination of the states we intend to estimate.
- We have an additional (inertia) condition that must be satisfied for the filter to exist.

Question: Why the *formal similarities* in the structure of the solutions, despite *fundamental differences* in the philosophies of the approaches?

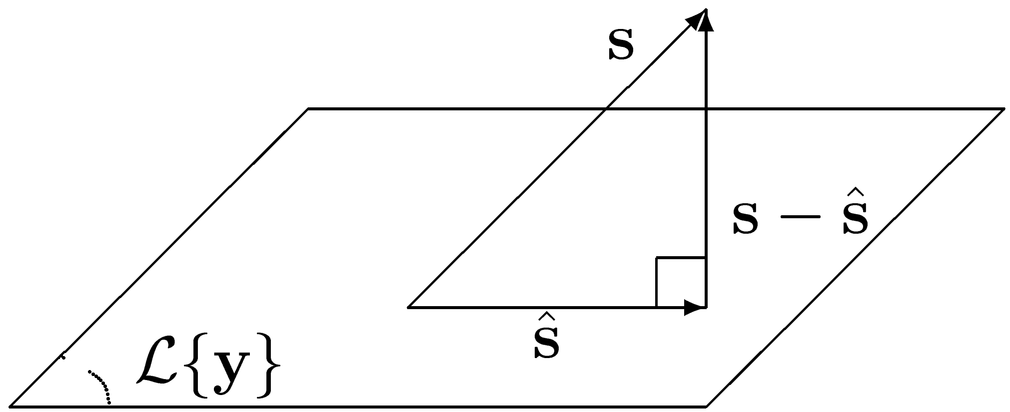
Consider the linear space (more precisely module), $\mathcal{P}$, of zero-mean vector-valued random processes with inner product

$$\langle \mathbf{x}_i, \mathbf{x}_2 \rangle \stackrel{\Delta}{=} E\mathbf{x}_1\mathbf{x}_2^*.$$

Then the *observations* space, $\mathcal{L}\{\mathbf{y}\}$, is defined as

$$\mathcal{L}\{\mathbf{y}\}\Delta{=}\mathcal{L}\{\mathbf{y}_0, \ldots, \mathbf{y}_N\} \quad , \quad \mathbf{y}^T = \left[\begin{array}{c} \mathbf{y}_0, \ldots \mathbf{y}_N \end{array} \right].$$

It is wellknown that $\hat{\mathbf{s}}$, the linear least-mean-square estimate of the random variable $\mathbf{s}$ is given by the projection onto $\mathcal{L}\{|bfy\}$.



- The solution to the (traditional) H^2 estimation problem is most naturally defined as a projection in a Hilbert space.
- When the observations have state-space structure, the Kalman filter is a method for recursively computing this projection.
- But what does this mean for H^∞ problems where there are no random variables?
- We have shown that H^∞ filters can also be obtained by a projection, not in a Hilbert space, but in a Krein space.
- The reason why H^∞ filters resemble Kalman filters is because they really are Kalman filters in Krein space.
- More generally, this simple observation explains both the (surprising) similarities and the differences between H^2 and H^∞ theories.
- The original solutions to the two theories were based on very different approaches.

Some of the original approaches to H^∞ theory include:

- interpolation-theoretic (Nehari problem)
- game-theoretic (dynamic programming)
- J-spectral factorization
- J-lossless conjugation
- bounded-real and positive-real lemmas
- mixed Hankel-plus-Toeplitz operators

Krein Spaces

Definition 1 *An abstract vector space $\{\mathcal{K}, \langle\cdot,\cdot\rangle\}$ that satisfies the following requirements is called a Krein space:*

1. *$\mathcal{K}$ is a linear space over $\mathcal{C}$, the complex numbers.*
2. *There exists a bilinear form $\langle\cdot,\cdot\rangle$ such that*
   (a) $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{x}\rangle^*$
   (b) $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{z}\rangle + b\langle \mathbf{y}, \mathbf{z}\rangle$ *for any* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{K}$, $a, b \in \mathcal{C}$*: $*$ denotes complex conjugation.*
3. *The vector space $\mathcal{K}$ admits a direct orthogonal sum decomposition*

$$\mathcal{K} = \mathcal{K}_+ \oplus \mathcal{K}_-$$

such that $\{\mathcal{K}_+, \langle\cdot,\cdot\rangle\}$ and $\{\mathcal{K}_-, -\langle\cdot,\cdot\rangle\}$ are Hilbert spaces and $\langle \mathbf{x}, \mathbf{y}\rangle == 0$ for any $\mathbf{x} \in \mathcal{K}_+$ and $\mathbf{y} \in \mathcal{K}_-$.

A Finite-Dimensional Krein Space

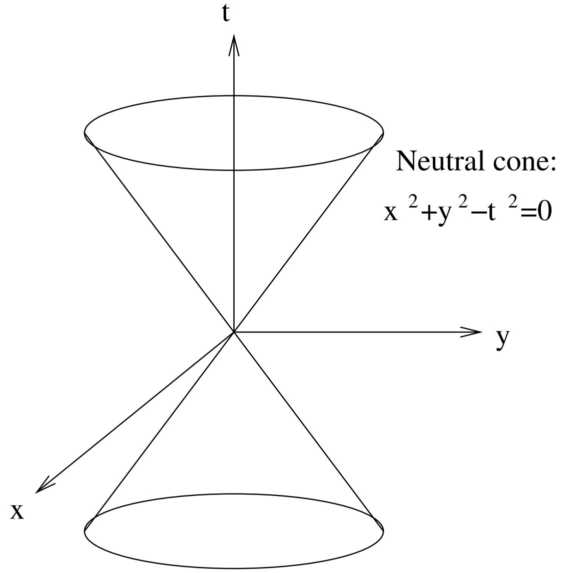


Figure 1: 3-dimensional Minkowski space

$$\mathbf{x} = \begin{bmatrix} x & y & t \end{bmatrix} \quad , \quad \langle \mathbf{x}, \mathbf{x} \rangle = x^2 + y^2 - t^2$$

Positive space $\quad \{\mathbf{x} | \langle \mathbf{x}, \mathbf{x} \rangle > 0\}$ (outside the cone)

Neutral space $\quad \{\mathbf{x} | \langle \mathbf{x}, \mathbf{x} \rangle = 0\}$ (cone)

Negative space $\quad \{\mathbf{x} | \langle \mathbf{x}, \mathbf{x} \rangle < 0\}$ (inside the cone)

Peculiarities of Krein Spaces

Although Krein spaces and Hilbert spaces share many characteristics, they differ in special ways that mark the differences between the H^∞ and H^2 theories. These include

- The existence of
  - *neutral* vectors: non-zero vectors with zero inner product
  - *isotropic* vectors: non-zero vectors in a linear subspace that are orthogonal to every vector in that subspace
- The non-existence or non-uniqueness of projections. Existence and uniqueness requires the nonsingularity of a certain Gramian matrix.
- One can only assert that quadratic forms in Krein space have stationary points. Further conditions must be met for a minimum.

Projections in Krein Space

Let us focus on projections.

Definition 2 *Given the elements* $\mathbf{z}$ *and* $\mathbf{y}_0, \ldots, \mathbf{y}_N$ *in* $\mathcal{K}$*, we define* $\hat{\mathbf{z}}$ *to be the projection of* $\mathbf{z}$ *onto* $\mathcal{L}\{\mathbf{y}\}$ *if* $\hat{\mathbf{z}} \in \mathcal{L}\{\mathbf{y}\}$ *and* $\mathbf{z} - \hat{\mathbf{z}}$ *satisfies the orthogonality condition,*

$$\mathbf{z} - \hat{\mathbf{z}} \perp \mathcal{L}\{\mathbf{y}\} \quad \textit{i.e.,} \quad \langle \mathbf{z} - \hat{\mathbf{z}}, \mathbf{y}_i \rangle = 0 \ \textit{ for all } i = 0, \ldots, N.$$

In Hilbert space, projections always exist and are unique. In Krein spaces, however,...

Theorem 1 *In a Krein space the projection exists and is unique iff, the Gramian matrix* $R_y = \langle \mathbf{y}, \mathbf{y} \rangle \stackrel{\Delta}{=} \left[\langle \mathbf{y}_i, \mathbf{y}_j \rangle_{i,j=0}^N \right]$ *is nonsingular. In this case, the projection is given by*

$$\hat{\mathbf{z}} = \langle \mathbf{z}, \mathbf{y} \rangle \langle \mathbf{y}, \mathbf{y} \rangle^{-1} \mathbf{y} = R_{zy} R_y^{-1} \mathbf{y},$$

where we have defined the cross-Gramian matrix $R_{zy} \stackrel{\Delta}{=} \left[\langle \mathbf{z}, \mathbf{y}_j \rangle_{i,j=0}^N \right]$.

Projections and Minimization

In Hilbert spaces, projections perform minimization. However, there is some confusion, because projections perform two *different* minimizations:

- a *stochastic* minimization

$$\arg \min_{\mathbf{z}_y \in \mathcal{L}\{\mathbf{y}\}} \langle \mathbf{z} - \mathbf{z}_y, \mathbf{z} - \mathbf{z}_y \rangle = \hat{\mathbf{z}} = R_{zy} R_y^{-1} \mathbf{y}.$$

- a *deterministic* minimization

$$\arg\min_z \begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} z \\ y \end{bmatrix} = \hat{z} = R_{zy} R_y^{-1} y.$$

Note that in the second problem z and y are not random variables, and so $\hat{z}$ is *not* the projection of z onto y. The point is that the solution is given by the same coefficient matrix that projects $\mathbf{z}$ onto $\mathbf{y}$.

In Krein spaces, however, projections, when they exist generally perform stationarization. For minimization, we need further conditions.

Theorem 2 *In Krein space, the projection* $\hat{\mathbf{z}} = R_{zy}R_y^{-1}\mathbf{y}$ *minimizes the stochatistic criterion*

$$\min_{\mathbf{z}_y \in \mathcal{L}\{\mathbf{y}\}} \langle \mathbf{z} - \mathbf{z}_y, \mathbf{z} - \mathbf{z}_y \rangle$$

iff,

$$R_y > 0.$$

It minimizes the deterministic criterion

$$\min_z \begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} z \\ y \end{bmatrix}$$

iff,

$$R_z - R_{zy}R_y^{-1}R_{yz} > 0.$$

Thus, although in Krein spaces projections stationarize both the stochastic and deterministic quadratic forms, they do not necessarily minimize both at the same time.

In Hilbert space, this is not an issue since

$$\left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix} > 0$$

which is equivalent to

$$R_y > 0 \quad \text{and} \quad R_z - R_{zy} R_y^{-1} R_{yz} > 0.$$

State-Space Structure and the Krein Space Kalman Filter

Theorem 3 *Consider the Krein state-space model*

$$\begin{cases} \mathbf{x}_{i+1} & F_i\mathbf{x}_i + G_i\mathbf{u}_i \\ \mathbf{y}_i & H_i\mathbf{x}_i + \mathbf{v}_i \end{cases} \qquad 0 \le i \le N$$

with

$$\left\langle \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \\ \mathbf{x}_0 \end{bmatrix}, \begin{bmatrix} \mathbf{u}_j \\ \mathbf{v}_j \\ \mathbf{x}_0 \end{bmatrix} \right\rangle = \begin{bmatrix} \begin{bmatrix} Q_i & 0 \\ 0 & R_i \end{bmatrix} \delta_{ij} & 0 \\ 0 & \Pi_0 \end{bmatrix}.$$

Then the projection of $s_i = L_i\mathbf{x}_i$ *onto* $\{\mathbf{y}_0, \ldots, \mathbf{y}_{i-1}\}$ *is given by* $\hat{\mathbf{s}}_i = L_i\hat{\mathbf{x}}_i$ *where*

$$\hat{\mathbf{x}}_{i+1} = F_i\hat{\mathbf{x}}_i + K_{p,i}(\mathbf{y}_i - H_i\hat{\mathbf{x}}_i), \quad \hat{\mathbf{x}}_0 = 0$$

where $K_{p,i} = F_i P_i H_i^* R_{e,i}^{-1}$, $R_{e,i} \stackrel{\Delta}{=} \langle \mathbf{y}_i - H_i \hat{\mathbf{x}}_i, \mathbf{y}_i - H_i \hat{\mathbf{x}}_i \rangle = R_i + H_i P_i H_i^*$ *and* P_i *satisfies the Riccati recursion*

$$P_{i+1} = F_i P_i F_i^* + G_i Q_i G_i^* - K_{p,i} R_{e,i} K_{p,i}^*, \quad P_0 = \Pi_0.$$

Moreover,

- *The Gramian matrix* R_y *is strongly regular, so that recursive projection exists, iff,* $R_{e,i}$ *is invertible for all* $0 \le i \le N$.
- *The stochastic problem has a minimum, i.e.,* $R_y > 0$, *iff* $R_{e,i} > 0$ *for all* $0 \le i \le N$.
- *The deterministic quadratic form that is stationarized by the projection is*

  $$x_0^* \Pi_0^{-1} x_0 + \sum_{i=0}^{N} u_i^* Q_i^{-1} u_i + \sum_{i=0}^{N} (y_i - H_i x_i)^* R_i^{-1} (y_i - H_i x_i).$$

  The condition $R_z - R_{zy} R_y^{-1} R_{yz} > 0$ *for a minimum is equivalent to* $R_{e,i}$ *and* R_i *having the same inertia for all* $0 \le i \le N$.

Back to H^∞ Estimation

What does all this have to do with H^∞ problems. Well, recall that the objective was to bound the energy gain by γ^2, i.e.,

$$\frac{\sum_{i=0}^{N}(\hat{s}_{i|i}-L_ix_i)^*(\hat{s}_{i|i}-L_ix_i)}{x_0^*\Pi_0^{-1}x_0+\sum_{i=0}^{N}u_i^*u_i+\sum_{i=0}^{N}(y_i-H_ix_i)^*(y_i-H_ix_i)}<\gamma^2,$$

for all disturbances x_0 and $\{u_i\}$. Now this is equivalent to

$$x_0^*\Pi_0^{-1}x_0+\sum_{i=0}^{N}u_i^*u_i+\sum_{i=0}^{N}(y_i-H_ix_i)^*(y_i-H_ix_i)-\gamma^{-2}\sum_{i=0}^{N}(\hat{s}_{i|i}-L_ix_i)^*(\hat{s}_{i|i}-L_ix_i),$$

which can be more compactly written as

$$x_0^*\Pi_0^{-1}x_0+\sum_{i=0}^{N}u_i^*u_i+\sum_{i=0}^{N}\begin{bmatrix} y_i-H_ix_i \\ \hat{s}_{i|i}-L_ix_i \end{bmatrix}^*\begin{bmatrix} I & 0 \\ 0 & -\gamma^2 \end{bmatrix}^{-1}\begin{bmatrix} y_i-H_ix_i \\ \hat{s}_{i|i}-L_ix_i \end{bmatrix}.$$

But this is simply the deterministic quadratic form correponding to the Krein space state-space model with parameters

$$Q_i \leftarrow I \quad , \quad R_i \leftarrow \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 \end{bmatrix} \quad , \quad H_i \leftarrow \begin{bmatrix} H_i \\ L_i \end{bmatrix}$$

The solution is straightforwardly given by the Krein space Kalman filter.

- The appearance of the indefinite covariance matrices (e.g., $\begin{bmatrix} I & 0 \\ 0 & -\gamma^2 \end{bmatrix}$) in the solution and Riccati recursion are now clear.
- The appearance of the L_i in the solution is also clear because it has entered the deterministic quadratic form.
- The inertia condition required for the existence of a solution is simply the conditon for the projection to minimize the quadratic form.

Linear Estimation in Krein Spaces

Krein space estimation yields a unified approach for studying and solving the following problems

- conventional H^2 estimation and control
- H^∞ estimation and control
- quadratic game theory
- risk-sensitive estimation and control
- finite memory adaptive filtering
- others

These (seemingly different) problems are all united in the sense that they can be related to certain *indefinite* quadratic forms from which a Krein space optimization problem can be inferred.

- H^∞ problems:

$$\frac{\|\cdot\|^2}{\|\cdot\|^2} < \gamma^2$$

$$\Rightarrow \ \|\cdot\|^2 - \|\cdot\|^2 > 0$$

- risk-sensitive problems:

$$\min E e^{\|\cdot\|^2}$$

$$\Rightarrow \ \min \int e^{\|\cdot\|^2} e^{-x^* R_x^{-1} x} = \min e^{\min_x \left[\|\cdot\|^2 - x^* R_x^{-1} x\right]}$$

- quadratic game problems

$$\min_u \max_w J(u, w, y), \quad J \text{ quadratic}$$

Apart from providing a unified derivation for various problems in estimation and control, this approach has the major bonus of allowing one to generalize to the H^∞ setting many of the wellknown results accumulated over the last four decades in conventional Kalman filtering and LQG theory.

These include:

- developing square-root and fast array algorithms for H^∞ problems
- Studying H^∞ adaptive filtering and demonstrating the H^∞-optimality of the LMS algorithm
- Giving a stochastic interpretation and geometric proof of the KYP (Kalman-Yakubovich-Popov) lemma
- Investigating the asymptotic behavior of H^∞ filters (and controllers)
- Studying the existence and properties of solutions to Riccati equations with (possibly) indefinite coefficient matrices

Conventional Square-Root Algorithm

We have seen that in the conventional Kalman filter

$$\hat{\mathbf{x}}_{i+1|i+1} = F_i \hat{\mathbf{x}}_{i|i} + K_{f,i+1}(y_{i+1} - H_{i+1}\hat{\mathbf{x}}_{i|i}), \quad \hat{\mathbf{x}}_{-1|-1}$$

the gain vector $K_{f,j}$ is updated using a certain Riccatu recursion. For various reasons (reduced dynamic range, better conditioning, stabler algorithms, etc.) attention has moved in the Kalman filtering community to so-called square-root array algorithms that propagate square-root factors of P_i.:

$$\begin{bmatrix} R_i^{1/2} & H_i P_i^{1/2} \\ 0 & P_i^{1/2} \end{bmatrix} \Theta_i^{(1)} = \begin{bmatrix} R_{e,i}^{1/2} & 0 \\ K_{f,i} R_{e,i}^{1/2} & P_{i|i}^{1/2} \end{bmatrix}$$

$$\begin{bmatrix} F_i P_{i|i}^{1/2} & G_i Q_i^{1/2} \end{bmatrix} \Theta_i^{(2)} = \begin{bmatrix} P_{i+1}^{1/2} & 0 \end{bmatrix}$$

where $\Theta_i^{(1)}$ and $\Theta_i^{(2)}$ are *any* unitary matrices that triangularize the pre-array. The algorithm is initialized with $P_0^{1/2} = \Pi_0^{1/2}$.

H^∞ Square-Root Algorithm

Developing H^∞ square-root algorithms from scratch is challenging. However, the Krein space connection show the way: square-roots are replaced by indefinite square-roots and unitary transformations by J-unitary ones.

H^∞ Square-Root Filter The level-*gamma* H^∞ filtering problem has a solution iff, for all $i = 0, \ldots, N$ there exist J-unitary (with $J = I \oplus -I \oplus I$) unitary matrices $\Theta_i^{(1)}$ (i.e., $\Theta_i^{(1)} J \Theta_i^{(1)*} = J$) such that

$$\begin{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \gamma I \end{bmatrix} & \begin{bmatrix} H_i \\ L_i \end{bmatrix} P_i^{1/2} \\ 0 & P_i^{1/2} \end{bmatrix} \Theta_i^{(1)} = \begin{bmatrix} R_{e,i}^{1/2} & 0 \\ K_{f,i} R_{e,i}^{1/2} & P_{i|i}^{1/2} \end{bmatrix}$$

$$\begin{bmatrix} F_i P_{i|i}^{1/2} & G_i Q_i^{1/2} \end{bmatrix} \Theta_i^{(2)} = \begin{bmatrix} P_{i+1}^{1/2} & 0 \end{bmatrix}$$

where $R_{e,i}^{1/2}$ and $P_{i+1}^{1/2}$ lower block triangular and with $\Theta_i^{(2)}$ unitary.

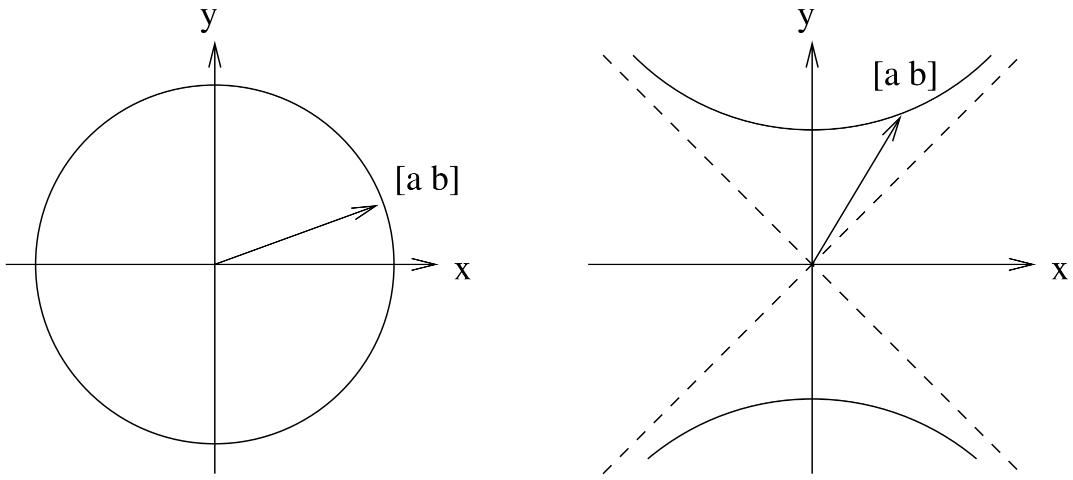
- The above algorithm has the interesting property that there is no need to check the inertia conditions required for the existence of a solution.
- In fact, the condition for the existence of the triangularizing J-unitary transformations is precisely the inertia conditions we need to check.
- In other words, if we can run the algorithm, then a solution exists and vice-versa.

- The point is that, unlike with unitary transformations where it is always possible to rotate vectors to lie in any direction, with J-unitary transformations this may or may not be possible.

  Take, for example, $J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and ask whether it is possible to find a J-unitary transformation Θ such that

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} c & 0 \end{bmatrix}.$$

  Clearly this can be done iff, $|a|^2 - |b|^2 \geq 0$.



Euclidean space   Minkowski Space

The Infinite Horizon Case

Now when the system matrices are constant, i.e., the state-space model is time-invariant

$$\begin{cases} \mathbf{x}_{i+1} = F\mathbf{x}_i + G\mathbf{u}_i \\ \mathbf{y}_i = H\mathbf{x}_i + \mathbf{v}_i \end{cases} \qquad i \geq 0$$

one may attempt to solve the H^2 or H^∞ problems in the infinite-horizon case.

- In this case it turns out that the optimal filters are also time-invariant.
- In the H^2 case, they can be obtained via spectral factorization
- In the H^∞ case they can be obtained via J-spectral factorization
- In both situations, the (J) spectral factorizations can be obtained by solving for the stabilizing solution of an (indefinite) *discrete-time algebraic Riccati equation* (DARE).

Convergence of the Riccati Recursion

In the time-invariant case, it is therefore natural to ask whether the solution to the Riccati recursion

$$P_{i+1} = FP_iF^* + GQG^* - FP_iH^*(R + HP_iH^*)^{-1}HP_iF^*, \quad P_0 = \Pi_0$$

converges to the corresponding stabilizing solution of the DARE

$$P = FPF^* + GQG^* - FPH^*(R + HPH^*)^{-1}HPF^*.$$

In the H^2 case, where $Q > 0$ and $R > 0$, it is well known that when $\{F, H\}$ is detectable and $\{F, GQ^{1/2}\}$ is stabilizable, then for all non-negative definite initial conditions $P_0 \geq 0$, the Riccati variable P_i converges to P.

However, extending a continuous-time result of Wonham, Kucera and Ljung and Kailath to the discrete-time case, one can show that convergence is guaranteed for all initial conditions that satisfy

$$I + (P^a)^{*/2} P_0 (P^a)^{1/2} > 0, \tag{1}$$

where P^a is the stabilizing solution to the *dual* DARE

$$P^a = F^* P^a F + H^* R^{-1} H - F^* P^a G (Q^{-1} + G^* P^a G)^{-1} G^* P^a F.$$

Note that this allows for some indefinite (and even negative semi-definite) initial conditions.

- But given that P_0 is the covariance matrix of the random initial state, what does this mean?

It turns out that (1) is equivalent to the condition that

$$R_{e,i} = R + HP_iP^* > 0, \quad \forall i \geq 0.$$

But this means that $R_y > 0$.

- In other words, we have convergence for all initial that lead to an output process $\{\mathbf{y}_i\}$ that has positive Gramian.
- To be more specific, it is not important that the internal model be physical, but that its output be so.

The above result can be generalized to Riccati recursions with indefinite covariance matrices to show that convergence occurs *if* $R_{e,i}$ *has constant inertia for all* $i \geq 0$.

But this, incidentally, is the condition for the existence of H^∞ filters, and thereby establishes their convergence.

The KYP Lemma

The KYP lemma has to do with whether the Popov function

$$\Sigma(z) = \begin{bmatrix} H(zI-F)^{-1} & I \end{bmatrix} \begin{bmatrix} GQG^* & GS \\ S^*G^* & R \end{bmatrix} \begin{bmatrix} (z^{-1}I-F^*)^{-1} \\ I \end{bmatrix},$$

is positive definite on the unit circle $|z| = 1$, or not.

When the center matrix is the covariance matrix of some random variables, i.e.,

$$\begin{bmatrix} GQG^* & GS \\ S^*G^* & R \end{bmatrix} \delta_{ij} = E \begin{bmatrix} Gu_i \\ v_i \end{bmatrix}^* \begin{bmatrix} Gu_j \\ v_j \end{bmatrix},$$

then the Popov function is a power spectral density function and it is clearly positive on the unit circle.

However, even if the center matrix is not positive definite, it is possible that the Popov function be so. Under an observability assumption on $\{F, H\}$, the KYP lemma asserts that this is the case iff, there exists a Hermitian Z such that

$$\begin{bmatrix} -Z + FZF^* + GQG^* & FZH^* + GS \\ HZF^* + S^*G^* & R + HZH^* \end{bmatrix} > 0.$$

But what does this mean?

It turns out that Z is the state covariance of a Krein space state-space model whose output is neutral, i.e,

$$\begin{bmatrix} H(zI - F)^{-1} & I \end{bmatrix} \begin{bmatrix} -Z + FZF^* & FZH^* \\ HZF^* & HZH^* \end{bmatrix} \begin{bmatrix} (z^{-1}I - F^*)^{-1} \\ I \end{bmatrix} = 0.$$

Therefore the KYP lemma states that if $\Sigma(z)$ is positive on the unit circle, then there is a positive-definite center matrix that generates it and that this can be obtained by adding a neutral output process to the original one.

Conclusion

- We have developed a unified approach to wide variety of estimation and control problems (LQG, H^∞, risk-sensitive, etc.) by considering estimation problems and control problems in an indefinite metric (Krein) space.
- The theory both explains the similarities and marks the differences (such as the difference between the H^2 and H^∞ separation principles) between the solutions to these different types of problems.
- It allows one to extend to these new settings many of the results obtained in the last four decades in Wiener-Kalman filtering and LQG control (such as numerically stable and computationally efficient algorithms)
- It also adds further insight into the classical problems (such as convergence for H^2 Riccati recursions, the KYP lemma, etc.).