The geometry of low-rank Kalman filters

S. Bonnabel (Mines ParisTech)
Joint work with
R. Sepulchre (Université de Liège)

CAS, Paris, 16 Fev 2012

Introduction: proof of concept

- **The natural metric of the cone of positive definite matrices** allows to analyse the convergence properties of the Kalman filter through the contraction properties of Riccati flow.
- We introduced in 2009[1] a new metric on the set of positive semi-definite matrices of fixed rank that generalizes the natural metric of the cone.
- *This metric is appropriate to analyse the contraction properties of a low-rank version of the Riccati flow.*

[1] Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. Bonnabel and Sepulchre. SIAM J. Matrix Anal. Appl. 2009.

Outline

1. Contraction properties of the Riccati flow
2. Low rank Kalman filtering
3. An invariant metric on the set of fixed rank psd matrices
4. Contraction properties of the low-rank Riccati flow

Kalman filter, discrete Riccati equation

System

$$\begin{aligned} x_t &= A_t x_{t-1} + B_t w_t \\ y_t &= C_t x_t + v_t, \;\; t = 0, 1, \ldots \end{aligned}$$

Kalman filter

$$\begin{aligned} \hat{x}_t &= (A_t - P_t C_t^T C_t)\hat{x}_{t-1} + P_t C_t^T y_t \\ P_t &= \Phi_t(P_{t-1}) \end{aligned}$$

Discrete Riccati Equation

$$\Phi_t(P) = (A_t P A_t^T + B_t B_t^T)(I + C_t^T C_t B_t B_t^T + C_t^T C_t A_t P A_t^T)^{-1}$$

Mapping Φ_t is positive on the cone $P_+(n)$.

Natural metric of $P_+(n)$

Group action of $GL(n)$ on $P_+(n)$ via congruence:

$$\varphi_A(P) = A^T P A$$

Reductive homogeneous space: $P_+(n)$ admits a $GL(n)$-invariant metric.

Metric based on Euclidian metric at Identity + invariance

$$g_P(D_1, D_2) = \mathrm{Tr}(P^{-1/2} D_1 P^{-1} D_2 P^{-1/2})$$

Geodesic distance: $d(P, Q) = \left(\sum_k \log^2(\lambda_k)\right)^{1/2}$ where $\det(PQ^{-1} - \lambda_k I) = 0$.

Invariance to congruence and inversion (=isometries)

$$d(P, Q) = d(A^T PA, A^T QA) = d(P^{-1}, Q^{-1})$$

Contraction of the Ricatti equation

Proposition [Bougerol (1993)] for $B \succ 0$ and $C \succ 0$, the map

$$\Phi(P) = (APA^T + B)(C(APA^T + B) + I)^{-1}$$
$$= (C + (APA^T + B)^{-1})^{-1}$$

is a contraction on the Riemmanian manifold $(P_+(n), g)$

$$\exists \gamma < 1 \quad \forall P_1, P_2 \in P_+(n) : d(\Phi(P_1), \Phi(P_2)) \leq \gamma d(P_1, P_2)$$

Proof idea: $\Phi(\cdot)$ decomposes as

$$(\cdot)^{-1} \circ (\cdot + C) \circ (\cdot)^{-1} \circ (A \cdot A^T) \circ (\cdot + A^{-1} B A^{-T})$$

Composition of additions and isometries !

Left to prove: $P \mapsto P + C$ with $C \succ 0$ is a contraction (logical).

In continuous time

Kalman-Bucy filter

$$\frac{d}{dt}\hat{x} = (A - PC'(HH')^{-1}C)\hat{x} + PC'(HH')^{-1}y \tag{1}$$

$$\frac{d}{dt}P = \Phi_t(P) = AP + PA' + GG' - PC'(HH')^{-1}CP \tag{2}$$

Contraction property in the sense of Lohmiller and Slotine for the natural metric $\|\delta P\|_g^2 = g_P(\delta P, \delta P) = \mathrm{Tr}(P^{-1}\delta P P^{-1}\delta P)$

$$\frac{d}{dt}\delta P = A\delta P + \delta P A' - \delta P C'(HH')^{-1}CP - PC'(HH')^{-1}C\delta P$$

$$\frac{d}{dt}\|\delta P\|_g^2 \leq -2\mathrm{Tr}((P^{-1}\delta P P^{-1}\delta P)GG'P^{-1}) \leq -2\lambda\|\delta P\|_g^2$$

with $\lambda = \mu/p_{\max}$ where μ is a lower bound on GG', and $p_{\max}$ is an upper bound on P.

Contraction of Kalman filter

- **Desired property of a filter**: exponential forgetting of the initial condition
- Rooted in invariance properties of the natural metric
- The same property holds for the continuous-time Riccati flow

Outline

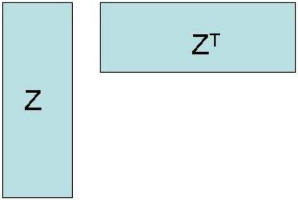
Low rank Kalman filter

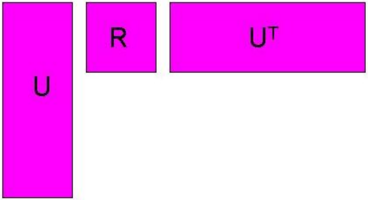
- Propagating the covariance matrix $P(t)$ is expensive in large-scale applications
- Choose $P(0)$ low-rank and propagate (a low-rank approximation of) $P(t)$ instead
- Quite popular in oceanography (e.g. SEEK Filter, Pham et al. 1997)
- Meets applications in EKF SLAM (see e.g. the SEIF)
- Natural metric is not defined on the facets. Geometry? Contraction?

Factorizations of P of rank $r \ll n$

Square-root factorization $P = ZZ^T$ with $Z \in \mathbb{R}_*^{n \times r}$: most popular, in particular in early numerical implementations of KF.

Polar factorization $P = URU^T$ with $(U, R) \in \mathrm{St(r,n)} \times \mathrm{P_+(r)}$: meaningful as U defines a subspace and R an ellipsoid.





Low rank Riccati equation

Riccati equation in polar factorization

$$\frac{d}{dt}URU' = AURU' + URU'A' + GG' - URU'C'(HH')^{-1}CURU'$$

which is not rank-preserving.

Modified rank-preserving Riccati equation: $GG' = \mu^2 UU'$

$$\frac{d}{dt}URU' = AURU' + URU'A' + \mu^2 UU' - URU'C'(HH')^{-1}CURU'$$

which is an alternative to the usual assumption $G = 0$.

Low rank Riccati equation

The update decomposes into the triangular system

$$\frac{d}{dt}U = (I - UU')AU \tag{3}$$

$$\frac{d}{dt}R = A_U R + R A_U + \mu^2 I - R C_U' (HH')^{-1} C_U R \tag{4}$$

with $A_U = U'AU$, $C_U = CU$

Subspace update (3) known as Oja flow which tracks the dominant subspace of A.

Low-rank cone update (4) is the projected Riccati equation.

Suggests to use a specific metric to study the contraction properties ?!

Outline

An invariant metric

Quotient manifold of psd matrices of fixed rank

$$S^+(r,n) \cong (\mathrm{St(r,n)} \times \mathrm{P_+(r)})/\mathrm{O(r)}$$

as $P = URU^T = (UO)(O^T RO)(O^T U^T)$ for any $O \in \mathrm{O(r)}$

Tangent space: $(\delta U, \delta R)$ tangent to $\mathrm{St(r,n)} \times \mathrm{P_+(r)}$ where $\delta U = U_\perp B$, i.e. is $\perp$ to the equivalence classes $\{UO\}$.

Proposed metric: sum of infinitesimal distances between subspaces and low-rank factors of $P_+(r)$

$$g((\delta U_1, \delta R_1), (\delta U_2, \delta R_2)) = \mathrm{Tr}(\delta U_1^T \delta U_2) + \mathrm{Tr}(R^{-1} \delta R_1 R^{-1} \delta R_2)$$

Invariant to rotations, scalings, pseudo-inversion.

Outline

Contraction properties

Oja flow contracting around the r-dominant subspace of $A^s = \frac{A+A'}{2}$ in the sense of Stiefel metric. Letting $\delta U = U_\perp B$

$$\frac{d}{dt}\mathrm{Tr}(\delta U'\delta U) = 2\mathrm{Tr}(BB'A^s_\perp - B'BA^s_r) \leq -\lambda_U \mathrm{Tr}(B'B) = -\lambda_U \mathrm{Tr}(\delta U'\delta U)$$

Low-rank cone flow contracting in the sense of natural metric

$$\frac{d}{dt}\mathrm{Tr}(R^{-1}\delta R R^{-1}\delta R) \leq -\lambda_R \frac{d}{dt}\mathrm{Tr}(R^{-1}\delta R R^{-1}\delta R)$$

Low rank Riccati equation contracting in the sense of the proposed metric around the dominant subspace of A^s

$$\begin{aligned}\frac{d}{dt} g((\delta U, \delta R), (\delta U, \delta R)) &= \frac{d}{dt}[\mathrm{Tr}(\delta U'\delta U) + \mathrm{Tr}(R^{-1}\delta R R^{-1}\delta R)] \\ &\leq -\lambda_U \mathrm{Tr}(\delta U'\delta U) - \lambda_R \mathrm{Tr}(R^{-1}\delta R R^{-1}\delta R)) \\ &\leq -\lambda g((\delta U, \delta R), (\delta U, \delta R))\end{aligned}$$

Contraction and convergence

$$\frac{d}{dt}U = (I - UU')AU \tag{5}$$

$$\frac{d}{dt}R = A_U R + R A_U + \mu^2 I - R C_U' (HH')^{-1} C_U R \tag{6}$$

For A symmetric $U \to U_r$. Thus $(A_U, C_U) \to (A_r, C_r)$, no peaking. Thus

The low-rank Riccati equation is eventually contracting in the sense of our metric with uniform rate.

For $A - A^T \neq 0$, U does not converge (can rotate forever). For topological reasons it can not contract everywhere.

Conclusion

- The Hilbert and natural metrics are useful to analyse contraction and convergence properties of positive maps
- Unfortunately they are not defined on the facets of $P_+(n)$
- One way to study the geometry of the facets is to fix the rank via $P = URU^T$. Our Riemannian geometry

$$S_+(r,n) \approx (St(r,n) \times P_+(r))/O(r)$$

  is insightful in several problems involving low-rank approximations of PSD matrices
- Here: it retains some of the contraction properties of the Riccati flow.
- Elsewhere: on $S_+(r,n)$ matrix means (filtering), online identification, optimization.