Optimal Sensor and Actuator Selection for Large-Scale Dynamical Systems

Neil K. Dhingra, Mihailo R. Jovanović, and Zhi-Quan Luo

UNIVERSITY OF MINNESOTA

49th Asilomar Conference on Signals, Systems and Computers, 2015

Motivation

- **Objective:** Systematic approach for sensor/actuator selection
- Applications
  - Phasor Measurement Units in power networks
  - Autonomous formations of vehicles
  - Flexible wing aircraft

Recent work

Sensor selection for parameter estimation

- Linear measurement model (Joshi, Boyd '09)
- Optimal experiment design (Kekatos, Giannakis, Wollenberg '12)

Sensor/actuator selection in dynamical systems

- Nonconvex formulation (Masazade, Fardad, Varshney '12)
- Adaptive selection (Chepuri, Leus '14)
- Maximize effect of actuators (Summers, Lygeros '14)

  (Tzoumas, Rahimian, Pappas, Jadbabaie '15)

Convex characterization as SDP

- SDP formulation (Polyak, Khlebnikov, Shcherbakov '13)
- Discrete time (Munz, Pfister, Wolfrum '14)

Sensor selection

Linear time invariant system with **many potential sensors**

$$\begin{aligned} \dot{x} &= A\,x + d \\ y &= C\,x + \eta \end{aligned}$$

Sensor selection

Linear time invariant system with **many potential sensors**

$$\begin{aligned}\dot{x} &= A\,x + d\\ y &= C\,x + \eta\end{aligned}$$

Kalman filter estimates state from measured output

$$\begin{aligned}\dot{\hat{x}} &= A\,\hat{x} + L\,(y - \hat{y})\\ \hat{y} &= C\,\hat{x}\end{aligned}$$

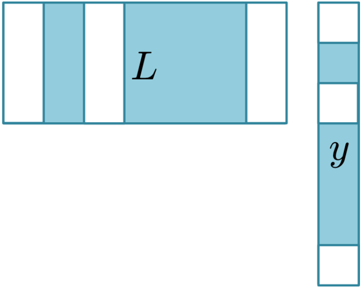
Sensor selection

Linear time invariant system with **many potential sensors**

$$\begin{aligned} \dot{x} &= A\,x + d \\ y &= C\,x + \eta \end{aligned}$$

Kalman filter estimates state from measured output

$$\begin{aligned} \dot{\hat{x}} &= A\,\hat{x} + L\,(y - \hat{y}) \\ \hat{y} &= C\,\hat{x} \end{aligned}$$



Objective: Minimize estimation error using a **few sensors**

Sensor selection via regularization

$$\text{minimize} \quad J(L) \quad + \quad \gamma \sum_i \| L \, \mathrm{e}_i \|_2$$

$J(L)$ ↓ **variance of estimation error**

$\gamma \sum_i \| L \, \mathrm{e}_i \|_2$ ↓ **column-sparsity-promoting penalty function**

- Larger γ selects fewer sensors
- $\gamma = 0$ uses all sensors

ℓ_1-penalized optimal control - Lin, Fardad, Jovanović, ACC '11, IEEE TAC '13

Observer design

$$\underbrace{\begin{aligned} \dot{x} &= A\,x + d \\ y &= C\,x + \eta \end{aligned}}_{\text{system}} \qquad \underbrace{\dot{\hat{x}} = A\,\hat{x} + L\,(y - \hat{y})}_{\text{observer}}$$

Variance of estimation error $e := (x - \hat{x})$

$$\begin{aligned} \dot{e} &= (A - LC)e + d + L\eta \\ J(L) &= \operatorname{trace}\left[\lim_{t\to\infty} \mathbb{E}\,\left(e(t)\,e^T(t)\right)\right] \end{aligned}$$

Kalman filter minimizes $J(L)$ using all sensors

SDP Characterization

V_d, V_η are process disturbance and measurement noise covariance

$$\underset{X,\, L}{\text{minimize}} \quad \underbrace{\text{trace}\left(V_d\, X\right) \; + \; \text{trace}(V_\eta\, L^T\, X\, L)}_{J(L)}$$

$$\text{subject to} \quad (A \, - \, L\, C)^T X \; + \; X\,(A \, - \, L\, C) \; + \; I \; = \; 0$$

$$X \; \succ \; 0$$

SDP Characterization

V_d, V_η are process disturbance and measurement noise covariance

$$\underset{X,\,L}{\text{minimize}} \quad \underbrace{\text{trace}\left(V_d\, X\right) \; + \; \text{trace}(V_\eta\, L^T\, X\, L)}_{J(L)}$$

$$\text{subject to} \quad (A \; - \; L\, C)^T\, X \; + \; X\, (A \; - \; L\, C) \; + \; I \; = \; 0$$

$$X \; \succ \; 0$$

Change of variables $Z := X\, L$

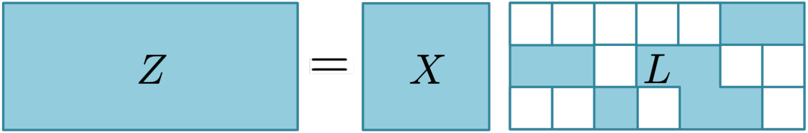
$$\underset{X,\,Z}{\text{minimize}} \quad \text{trace}\left(V_d\, X\right) \; + \; \text{trace}(V_\eta\, Z^T\, X^{-1}\, Z)$$

$$\text{subject to} \quad A^T\, X \; - \; C^T\, Z^T \; + \; X\, A \; - \; Z\, C \; + \; I \; = \; 0$$

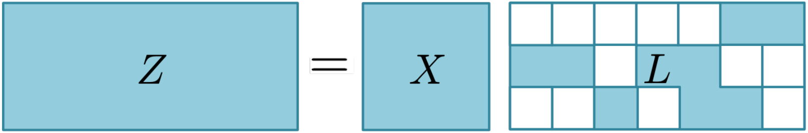
$$X \; \succ \; 0$$

- Schur complement yields convex problem

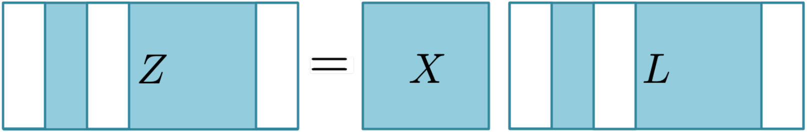
- Challenge: impose structure on L in (X, Z) coordinates
  - Linear constraints on L become *nonlinear* on (X, Z)



- Challenge: impose structure on L in (X, Z) coordinates
  - Linear constraints on L become *nonlinear* on (X, Z)



- Column-sparsity is preserved



Sensor selection - Semidefinite Program

$$\underset{X,\,Z}{\text{minimize}} \quad \underbrace{\text{trace}\left(V_d\, X\right) \; + \; \text{trace}(V_\eta\, Z^T\, X^{-1}\, Z\big) \; + \; \gamma \sum \|Z\, \mathrm{e}_i\|_2}_{f(X,\,Z)}$$

$$\text{subject to} \quad A^T X \; - \; C^T Z^T \; + \; X A \; - \; Z C \; + \; I \; = \; 0$$

$$X \; \succ \; 0$$

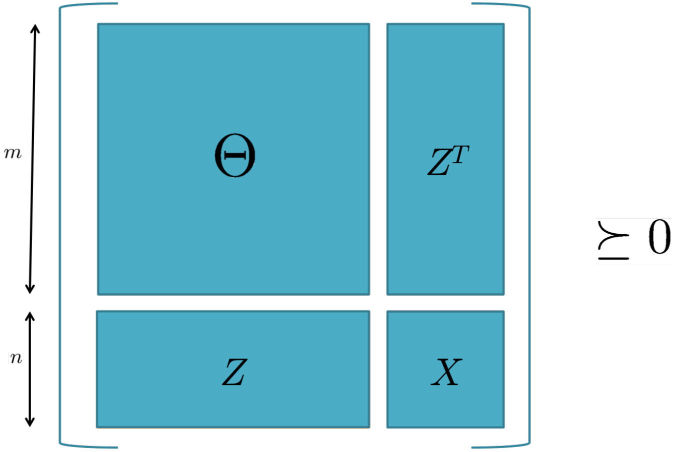
- Promote column sparsity of Z instead of L

Polyak, Khlebnikov, Shcherbakov, ECC '13

Dhingra, Jovanović, Luo, CDC '14

Computational complexity

$$\text{trace}(V_\eta Z^T X^{-1} Z) = \text{trace}(V_\eta \Theta)$$



Worst case computational complexity: $\mathcal{O}\left((n+m)^6\right)$

Alternating Direction Method of Multipliers (ADMM)

- **Splitting method** for convex problems with linear constraints

Form **augmented Lagrangian**

$$\mathcal{L}_\rho(X, Z, \Lambda) \;:=\; f(X, Z) \,+\, \langle \Lambda, h(X, Z) \rangle \,+\, \frac{\rho}{2}\,\|h(X, Z)\|_F^2$$

$$h(X, Z) \;:=\; A^T X \,-\, C^T Z^T \,+\, X A \,-\, Z C \,+\, I$$

- Update variables **separately**

$$
\begin{aligned}
X^{k+1} &= \arg\min_X \; \mathcal{L}_\rho(X,\, Z^k,\, \Lambda^k) \\
Z^{k+1} &= \arg\min_Z \; \mathcal{L}_\rho(X^{k+1},\, Z,\, \Lambda^k) \\
\Lambda^{k+1} &= \Lambda^k + \rho\, h(X^{k+1},\, Z^{k+1})
\end{aligned}
$$

Boyd, Parikh, Chu, Peleato, Eckstein ‘11

Z-minimization

$$\underset{Z}{\text{minimize}} \;\; \gamma \sum \|Z\,\mathrm{e}_i\|_2 \; + \; \underbrace{\frac{\rho}{2}\,\|Z\,C \; + \; C^T\,Z^T \; + \; W^k\|_F^2}_{J_Z(Z)}$$

- **Group LASSO**

Z-minimization

$$\underset{Z}{\text{minimize}} \;\; \gamma \sum \| Z \, \mathrm{e}_i \|_2 \; + \; \underbrace{\frac{\rho}{2} \| Z \, C \; + \; C^T \, Z^T \; + \; W^k \|_F^2}_{J_Z(Z)}$$

- **Group LASSO**

- Proximal method: iterative soft thresholding algorithm (ISTA)

$$Z^{m+1} \; = \; \mathcal{S}_{\alpha^m \gamma / \rho} \left(Z^m \; - \alpha^m \nabla J_Z(Z^m) \right)$$

- Computational complexity $\mathcal{O}(n^2 \, m)$

X-minimization

$$\underset{X}{\text{minimize}}\;\; \text{trace}\left(X\,V_d \;+\; X^{-1}\,Z^k\,V_\eta\,(Z^k)^T\right) + \frac{\rho}{2}\,\|A^T\,X + X\,A + U_k\|_F^2$$

- Can formulate as SDP (worst case $\mathcal{O}(n^6)$)
  - Many sensors, $m \gg n \implies n^6 \ll (n+m)^6$
  - Model reduction can reduce n

X-minimization

$$\underset{X}{\text{minimize}}\ \ \text{trace}\left(X\,V_d\ +\ X^{-1}\,Z^k\,V_\eta\,(Z^k)^T\right) + \frac{\rho}{2}\,\|A^T\,X + X\,A + U_k\|_F^2$$

- Can formulate as SDP (worst case $\mathcal{O}(n^6)$)
  - Many sensors, $m \gg n \implies n^6 \ll (n+m)^6$
  - Model reduction can reduce n

- Projected Newton's method
  - Conjugate gradient (worst case $\mathcal{O}(n^5)$)
  - Inexact minimization; faster in practice
  - Project onto $\{X : X \succ 0\}$

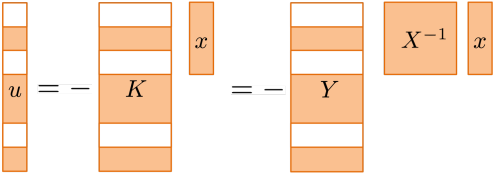
Extensions - Actuator selection

State feedback $u = -Kx$

$$\dot{x} = Ax + Bu + d$$

$$J(K) = \int_0^\infty x^T Q x + u^T R u \, \mathrm{d}t$$

- Promote row sparsity of $Y := KX$



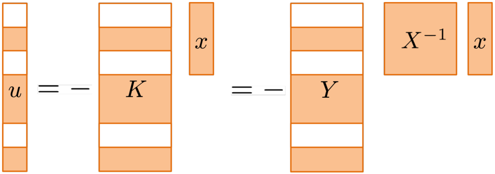
Extensions - Actuator selection

State feedback $u = -Kx$

$$\dot{x} = Ax + Bu + d$$

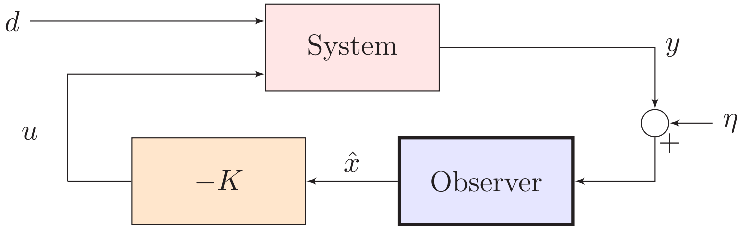
$$J(K) = \int_0^\infty x^T Q x + u^T R u \, \mathrm{d}t$$

- Promote row sparsity of $Y := KX$



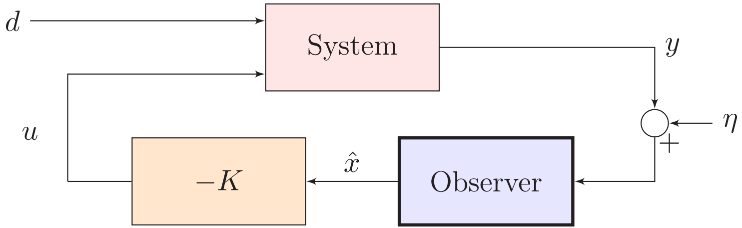
$$\begin{array}{ll} \underset{X,Y}{\text{minimize}} & \text{trace}(Q\,X) + \text{trace}(R\,Y\,X^{-1}\,Y^T) + \gamma \sum \|\mathrm{e}_i^T Y\|_2 \\ \text{subject to} & A\,X - B\,Y + X\,A^T - Y^T B^T + V_d = 0 \\ & X \succ 0 \end{array}$$

Extensions: Sensor selection for feedback control



- Given linear quadratic regulator $u = -K\hat{x}$
- Select sensors to optimize closed loop performance

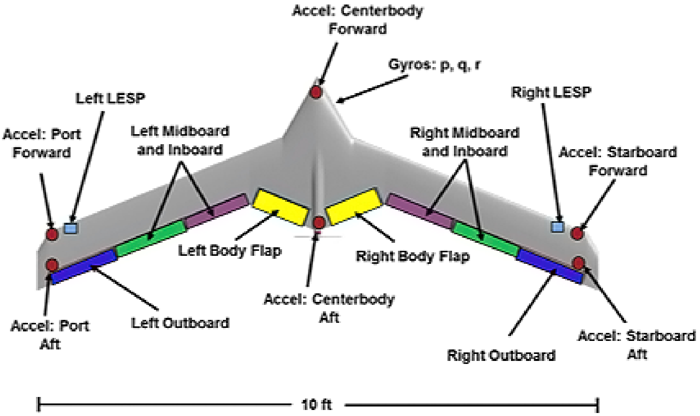
Extensions: Sensor selection for feedback control



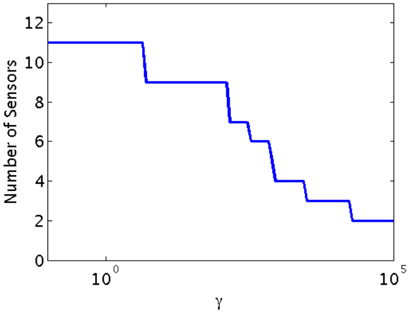
- Given linear quadratic regulator $u = -K\hat{x}$
- Select sensors to optimize closed loop performance

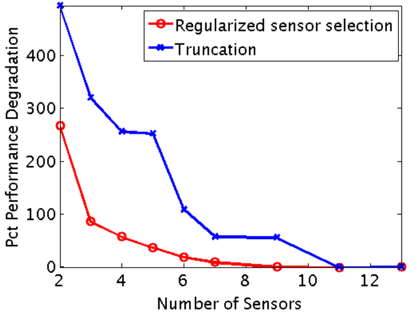
$$\begin{array}{ll} \underset{X,\,Z}{\text{minimize}} & \text{trace}\left(V_d\, X\right) \;+\; \text{trace}(V_\eta\, Z^T\, X^{-1}\, Z) \;+\; \gamma \sum \|Z\, \mathrm{e}_i\|_2 \\ \text{subject to} & A^T X \;-\; C^T Z \;+\; X\, A \;-\; Z\, C \;+\; {\color{red}K^T R K} \;=\; 0 \\ & X \;\succ\; 0 \end{array}$$

Example - Flexible Wing Aircraft



- Detect aeroelastic instability





- Using less than half the sensors degrades performance by only $\sim 20\%$

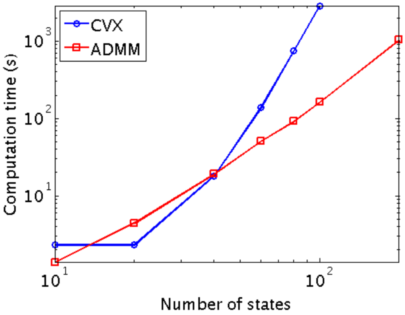
Example - Mass Spring System

- Dynamics

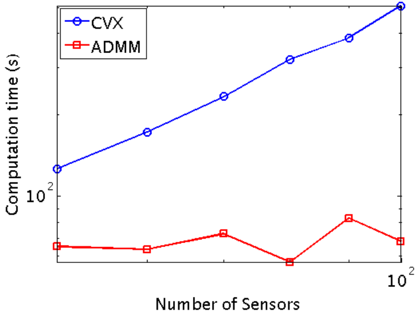
$$\ddot{x}_i = (x_{i+1} - x_i) + (x_{i-1} - x_i) + d_i$$

- Sensors
  - Position of each mass
  - Velocity of each mass

Computation time for $\gamma = 100$



n states and n outputs



50 states and m outputs

Conclusions

- Convex characterization of sensor or actuator selection
- Splitting algorithm
  - Solve simpler sub-problems
  - Scaling dependent on number of states

Future work

- More efficient X-minimization
- Alternative splitting methods
- Joint sensor and actuator selection

Acknowledgements

Support:

- NASA JPFP Fellowship
- MnDRIVE Graduate Scholars Program
- NSF ECCS-1407958

Collaboration on aircraft example

- Marty Brenner, NASA
- Claudia Moreno and Harald Pfifer

Dual Problem

$$g(Y) = \underset{e_i^T U \le \gamma}{\text{maximize}} \langle U, Y \rangle$$

Lagrangian

$$\begin{aligned} L(X, Y, \Lambda, M) &= \operatorname{trace}(QX + X^{-1}Y^T RY) + \langle U, Y \rangle \\ &\quad + \left\langle \Lambda, AX + XA^T - BY - Y^T B^T + V \right\rangle - \langle M, X \rangle \end{aligned}$$

where $M \succeq 0$

$$\begin{bmatrix} \nabla_X L \\ \nabla_Y L \end{bmatrix} = \begin{bmatrix} Q - X^{-1}Y^T RYX^{-1} + A^T\Lambda + \Lambda A \\ 2RYX^{-1} + U - 2B_2^T \Lambda \end{bmatrix}$$

Dual Problem

$$\begin{array}{ll} \text{maximize} & \langle V, \Lambda \rangle \\ \text{subject to} & Q + \Lambda A + A^T \Lambda - (B_2^T \Lambda - \frac{1}{2}U)^T R^{-1} (B_2^T \Lambda - \frac{1}{2}U) \succeq 0 \\ & \|e_i^T U\|_2 \le \gamma \end{array}$$