

# SENSAI FISHER INFORMATION, $K$ -RANK APPROXIMATIONS AND PARAMETER ESTIMATION

SIMON TAVENER

Date & time: 2016-10-23 17:49

**1. Introduction.** We seek  $\mathbf{x}(t, \mathbf{p})$  where  $\mathbf{x} \in \mathbb{R}^{\text{xdim}}$  and  $\mathbf{p} \in \mathbb{R}^{\text{kdim}}$  such that

$$\left. \begin{aligned} \frac{d\mathbf{x}}{dt} &= \mathbf{f}(t, \mathbf{x}, \mathbf{p}), \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{aligned} \right\} \quad (1.1)$$

and consider quantities of interest that are functions of both the solution and the parameters, i.e.,

$$\mathbf{q}(t) = \mathbf{q}(t, \mathbf{x}(t, \mathbf{p}), \mathbf{p}) \quad (1.2)$$

where  $\mathbf{q} \in \mathbb{R}^{\text{qdim}}$ .

The Matlab package SENSAl computes sensitivities and elasticities of both the solution variables  $\mathbf{x}$  and specified quantities of interest  $\mathbf{q}$ , with respect to parameters and initial conditions. It does so by constructing and solving a linear system of differential equations for the sensitivities (which are then scaled to provide elasticities). SENSAl utilizes MuPAD to perform symbolic differentiation to construct the Jacobian matrix

$$\mathbf{J} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \in \mathbb{R}^{\text{xdim} \times \text{xdim}} \quad (1.3)$$

and the matrix of partial derivatives with respect to the parameters

$$\mathbf{K} = \frac{\partial \mathbf{f}}{\partial \mathbf{p}} \in \mathbb{R}^{\text{xdim} \times \text{kdim}}, \quad (1.4)$$

and to write the Matlab routines to compute these derivatives.

**2. Fisher information matrix.** The sensitivity matrix  $\mathbf{M}$  defines the rate of change of the solution, or quantities of interest derived from the solution, with respect to the parameters. In general, the sensitivity matrix is defined as

$$\mathbf{M} = \frac{\partial \mathbf{q}}{\partial \mathbf{p}} \in \mathbb{R}^{\text{Mdim} \times \text{Fdim}}. \quad (2.1)$$

where

$$\text{Mdim} = \begin{cases} \text{xdim} & \text{if using primitive variables} \\ \text{qdim} & \text{if using quantities of interest} \end{cases} \quad (2.2)$$

and  $\text{Fdim}$  is the number of parameters we chose to vary. In SENSAl, their indices are contained in the vector  $Fp(1 : \text{Fdim})$ . The Fisher Information Matrix

$$\mathbf{F} = (\mathbf{M}^\top \mathbf{M})^{-1} \in \mathbb{R}^{\text{Fdim} \times \text{Fdim}}. \quad (2.3)$$

In practice, we never construct the  $\mathbf{F}$  but perform all operations on the sensitivity matrix  $\mathbf{M}$ . Consider the singular value decomposition

$$\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^\top, \quad \mathbf{U} \in \mathbb{R}^{\text{Mdim} \times \text{Mdim}}, \mathbf{S} \in \mathbb{R}^{\text{Mdim} \times \text{Fdim}}, \mathbf{V} \in \mathbb{R}^{\text{Fdim} \times \text{Fdim}} \quad (2.4)$$

We need to consider situations in which  $\text{Mdim} > \text{Fdim}$  and  $\text{Mdim} < \text{Fdim}$ . We define  $\text{Sdim}$ , the maximal rank of  $\mathbf{S}$ , as

$$\text{Sdim} := \min\{\text{Mdim}, \text{Fdim}\} \quad (2.5)$$

This is the maximum number of non-zero singular values of  $\mathbf{M}$ .

A simple calculation shows that the inverse of the Fisher Information matrix is unitarily similar to the  $\text{Sdim} \times \text{Sdim}$  diagonal matrix  $\mathbf{S}^\top \mathbf{S}$  and the non-zero real eigenvalues of the Fisher Information Matrix,  $\mathbf{F}$ , typically labelled  $\sigma_i, i = 1, \dots, \text{Sdim}$  are related to the singular values of the sensitivity matrix  $\mathbf{M}$  as

$$\sigma_i = \frac{1}{s_i^2}, \quad i = 1, \dots, \text{Sdim}. \quad (2.6)$$

There are two very different types of Fisher information matrix, one that is a (continuous) function of time discussed in §2.1 and one that is a function of a discrete times, discussed in §2.2.

**2.1. Continuous Fisher information matrix (gtype\_FIM).** The Fisher information matrix is based on the sensitivities or elasticities of the variables or quantities of interest *at all* times at which the solution is computed, i.e., at all times in the vector  $\mathbf{t}$ . The sensitivity matrix  $\mathbf{M}(t)$  is of size  $\text{Mdim} \times \text{Fdim}$ .

**2.2. Global Fisher information matrix (gtype\_FIM\_global).** A *globally* defined Fisher information matrix can also be computed based on the sensitivities or elasticities of the variables or quantities of interest *at all* times at which the solution is computed, i.e., at all times in the vector  $\mathbf{t}$ . The sensitivity matrix  $\mathbf{M}$  is now of size  $(\text{Mdim} * \text{ntsteps}) \times \text{Fdim}$ .

SENSAI also computes the parameter sensitivity score (pss) defined by [5].

**3. Active subspaces (gtype\_active\_subspace).** The Active Subspace method computes the singular value decomposition of the sensitivity matrix at all times at which the solution is computed. It further computes and plots the singular values as a function of time, and the singular vectors as a function of time.

**3.1. Local  $k$ -rank approximation in parameter space (surrogate model).** We also compute a low dimensional approximation to the function surface in parameter space around each solution point. There are at most  $\text{Sdim}$  non-zero singular values. Let  $\text{Adim} \leq \text{Sdim}$  be the dimension of the active subspace. Consider a linear approximation in parameter space about  $\mathbf{p}_0$  for either the solution  $\mathbf{x}$  or a vector-valued quantity of interest  $\mathbf{q}$ . Writing either as  $\mathbf{q}$ , we have

$$\mathbf{q}(t, \mathbf{p}_0 + \Delta \mathbf{p}) \approx \mathbf{q}(t, \mathbf{p}_0) + \left. \frac{\partial \mathbf{q}}{\partial \mathbf{p}} \right|_{\mathbf{q}(t, \mathbf{p}_0)} \Delta \mathbf{p} = \mathbf{q}(t, \mathbf{p}_0) + \mathbf{M} \Delta \mathbf{p} \quad (3.1)$$

where, as before

$$\mathbf{M} \in \mathbb{R}^{\text{Mdim} \times \text{Fdim}}. \quad (3.2)$$

Now

$$\begin{aligned} M\Delta p &= USV^\top \Delta p \\ &= \left( \sum_{i=1}^{Sdim} s_i \mathbf{u}_i \mathbf{v}_i^\top \right) \Delta p \end{aligned} \quad (3.3)$$

where  $\mathbf{u}_i$  is the  $i$ th column of  $U$ ,  $\mathbf{u}_i \in \mathbb{R}^{Mdim \times 1}$ , and  $\mathbf{v}_i^\top$  is the  $i$ th row of  $V^\top$ , i.e., the  $i$ th column of  $V$ ,  $\mathbf{v}_i^\top \in \mathbb{R}^{1 \times Fdim}$ . The best  $Adim$ -rank approximation of  $M\Delta p$  is

$$\mathbf{q}(t, \mathbf{p}_0 + \Delta \mathbf{p}) \approx \mathbf{q}(t, \mathbf{p}_0) + \left( \sum_{i=1}^{Adim} s_i \mathbf{u}_i \mathbf{v}_i^\top \right) \Delta \mathbf{p}. \quad (3.4)$$

**3.2. Angles between active subspaces.** We compute the angles between the initial singular vector(s) and the singular vector(s) at subsequent times using the work of [12]. Once the angle exceeds a critical value, the QoIs are deemed “geometrically distinct”.

### 3.3. Future work. *Ideas*

1. Set sample times once the angle between leading subspaces increases beyond a critical value.
2. Choose sample times at equally spaced angles between the leading subspaces.
3. Optimal sampling: Locate  $n$  sample points so as to maximize the angles between active subspaces. See [3]
4. Use adjoint techniques to compute  $M$  – see [6].

**4. Efficient forward funding evaluation for SIP (gtype\_SIP).** Algorithm to sample the  $kdim$ -dimensional parameter space

1. Locate `nsample` equally-spaced (global) nodes in the  $kdim$ -dimensional parameter space (squares).
2. Evaluate the sensitivity matrix and its singular value decomposition to determine the active subspace at each of these nodes.
3. Randomly sample the  $kdim$ -dimensional parameter space at `nsim` points.
4. Find the node closest to every sample point as measured *using the local FIM metric at each node*. The set of samples closest to each node in the local FIM metric is the Voronoi cell for that node as measured in the FIM metric.
5. Approximate the function using (3.4) at `msample` equally-spaced points in the active subspace local to each node. Call these local active subspace nodes
6. Project each sample point in to the active subspace within its Voronoi cell and determine the closest global or local active subspace node to the projection. (This is an inherently parallel computation.)
7. Assign the function value at the nearest node to the random sample.

**5. Parameter estimation (gtype\_GESS).** Given either the solution vector or quantity (quantities) of interest at all times for which the ode solution is known, SENSAL computes the generalized least squares best fit for the parameters.

### 6. Optimal experimental design. [1, 2, 5, 4, 5, 7, 10, 11]

*Idea:* Base optimal sampling strategies on the *angles between active subspaces*. Is this approach robust wrt error?. Other optimal sampling techniques based on the *spectrum* of the Fisher Information Matrix are highly sensitive to error.

*Question* What can we do if the QoI is not a differentiable function of the parameters, e.g., max voltage for the Hodgkin Huxley system of odes.

**7. Inverse sensitivities.** Create data files for SIP using SENSAL.

*Ideas*

1. Approximate function near sample points using the best  $k$ -rank approximation given by (3.4)
2. Optimal sampling: Locate  $n$  sample points so as to maximize the angles between active subspaces. See [3].
3. Find  $\mathbf{M}$  by solving an adjoint problem for a given scalar-valued quantity of interest – see [9] What can we do for a vector-valued quantities of interest? I believe we will need to solve multiple adjoint problems.
4. [8] uses adjoint ideas to produce simplified models, but do not seem to use the svd. Check this out.

**8. Examples.**

**8.1. Logistic growth.** Let

$$\mathbf{x} = x \in \mathbb{R}^1, \quad \mathbf{p} = (r, K)^\top \in \mathbb{R}^2. \quad (8.1)$$

Find  $x(t)$  such that

$$\begin{aligned} \frac{dx}{dt} &= rx \left(1 - \frac{x}{K}\right), \\ x(0) &= x_0. \end{aligned} \quad (8.2)$$

Quantities of interest,  $q(iq), iq = 1, \dots, \mathbf{qdim}$  are given by

$$q(1) = x, \quad q(2) = \frac{x}{K}, \quad q(3) = \frac{x}{x_0}. \quad (8.3)$$

**8.2. SEIR.** Let

$$\mathbf{x} = (S, E, I, R)^\top \in \mathbb{R}^4, \quad \mathbf{p} = (N, L, D, M, P, \beta_0, a, b)^\top \in \mathbb{R}^8. \quad (8.4)$$

Find

$$\begin{aligned} \frac{dS}{dt} &= \frac{N}{P} + \frac{R}{L} - \frac{\beta(t)SI}{N} - \frac{S}{P}, \\ \frac{dE}{dt} &= \frac{\beta SI}{N} - \frac{E}{M} - \frac{E}{P}, \\ \frac{dI}{dt} &= \frac{E}{M} - \frac{I}{D} - \frac{I}{P}, \\ \frac{dR}{dt} &= \frac{I}{D} - \frac{R}{L} - \frac{R}{P}, \end{aligned} \quad (8.5)$$

where

$$\beta(t) = \beta_0(1 + a \cos(2\pi t) + b \sin(2\pi t)) \quad (8.6)$$

with initial conditions

$$(S, E, I, R)^\top(0) = (S_0, E_0, I_0, R_0)^\top. \quad (8.7)$$

We choose quantities of interest

$$q(1) = \frac{I}{S + E + I + R}, \quad q(2) = \frac{E}{S + E + I + R}, \quad q(3) = \frac{R}{S + E + I + R}. \quad (8.8)$$

### 8.3. MSEIR. Let

$$\mathbf{x} = (M, S, E, I, R)^\top \in \mathbb{R}^5, \quad \mathbf{p} = (B, \delta, \mu_M, \beta, \mu_G, 1/\epsilon, \mu_I, 1/\gamma, f, \iota)^\top \in \mathbb{R}^{10}. \quad (8.9)$$

where  $M$  is the class protected by maternal antibodies. Find

$$\begin{aligned} \frac{dM}{dt} &= B(S + E + I + R) - (\delta + \mu_M)M \\ \frac{dS}{dt} &= \delta M - \beta SI - (\mu_G + \iota)S + fR \\ \frac{dE}{dt} &= \beta SI - (\epsilon + \mu_G)E \\ \frac{dI}{dt} &= \epsilon E - (\gamma + \mu_I + \mu_G)I \\ \frac{dR}{dt} &= \gamma I - (\mu_G + f)R + \iota S \end{aligned} \quad (8.10)$$

B	avg. birth rate	$[2.72E - 4, 3.04E - 4]$	$3.02E - 4$
$\delta$	avg. temporary immunity	$[1/12, 1/4]$	0.16
$\mu_M$	avg. infant death rate	$[4E - 3, 6E - 3]$	$4.5E - 3$
$\beta$	contact/infectivity rate	$[1.92E - 3, 3.85E - 3]$	$3.4E - 3$
$\mu_G$	avg. general death rate	$[2.4E - 4, 2.72E - 4]$	$2.52E - 4$
$1/\epsilon$	avg. infection time	$[0.571, 1]$	0.7
$\mu_I$	avg. infected death rate	$[4.81E - 6, 2.11E - 5]$	$1.75E - 5$
$1/\gamma$	avg. recovery time	$[0.7, 2.33]$	0.8
$f$	avg. loss of immunity rate	$[0.125, 0.25]$	0.18
$\iota$	avg. immunization rate	$[0.015, 0.0375]$	0.026
$M_0$	initial infants	$[2.5, 3.5]$	3.25
$S_0$	initial susceptibles	$[260, 275]$	270
$E_0$	initial exposed	$[0.01, 0.5]$	0.425
$I_0$	initial infected	$[0.1, 4]$	3.8
$R_0$	initial recovered/immunized	$[10, 20]$	13

TABLE 8.1

Uncertain parameters and initial conditions and their interval bounds in the MSEIRS model given by Eq. 8.10. The last column of values are the reference parameter values used to define the mean of the output densities.

### 8.4. Hodgkin Huxley. Let

$$\mathbf{x} = (V, m, h, n)^\top \in \mathbb{R}^4, \quad \mathbf{p} = (gNa, gK, gL)^\top \in \mathbb{R}^3. \quad (8.11)$$

Find

$$\begin{aligned} \frac{dV}{dt} &= II - gNa * m^3 * h * (V - vNa) - gK * n^4 * (V - vK) - gL * (V - vL), \\ \frac{dm}{dt} &= (1 - m) * aM - m * bM, \\ \frac{dh}{dt} &= (1 - h) * aH - h * bH, \\ \frac{dn}{dt} &= (1 - n) * aN - n * bN, \end{aligned} \quad (8.12)$$

with initial conditions

$$(V, m, h, n)^T(0) = (V_0, m_0, h_0, n_0)^T \quad (8.13)$$

where

$$\begin{aligned} II &= 20, \quad vNa = 115, \quad vK = -12, \quad vL = 10.6, \\ aM &= 0.1 * (25 - V) / (\exp((25 - V)/10) - 1), \\ aH &= 0.07 * \exp(-V/20), \\ aN &= 0.01 * (10 - V) / (\exp((10 - V)/10) - 1), \\ bM &= 4 * \exp(-V/18), \\ bH &= 1 / (\exp((30 - V)/10) + 1), \\ bN &= 0.125 * \exp(-V/80). \end{aligned} \quad (8.14)$$

We choose quantities of interest

$$q(1) = V, \quad q(2) = \frac{V}{m}, \quad q(3) = \frac{V}{h}, \quad q(4) = \frac{V}{n}. \quad (8.15)$$

#### REFERENCES

- [1] S.P. ASPREY AND S. MACCHIETTO, *Designing robust optimal dynamical experiments*, Journal of Process Control, 12 (2002), pp. 545–556.
- [2] H.T. BANKS, K. HOLM, AND F. KAPPEL, *Comparison of optimal design methods in inverse problems*, Inverse Problems, 27 (2011), pp. 165 – 178.
- [3] ÅKE BJÖRCK AND GENE H. GOLUB, *Numerical methods for computing angles between linear subspaces*, Mathematics of Computation, 27 (1973), pp. 579–594.
- [4] YUNFEI CHU AND JUERGEN HAHM, *Integrating parameter selection with experimental design under uncertainty for nonlinear dynamic systems*, AIChE Journal, 54 (2008), pp. 2310–2320.
- [5] A. CINTRÓN-ARIAS, H.T. BANKS, A. CAPALDI, AND A.L. LLOYD, *A sensitivity matrix based methodology for inverse problem formulation*, Journal of Inverse and Ill-Posed Problems, 17 (2009), pp. 545–564.
- [6] D. ESTEP AND D. NECKELS, *Fast and reliable methods for determining the evolution of uncertain parameters in differential equations*, Journal of Computational Physics, 213 (2006), pp. 530–556.
- [7] VALERII FEDEROV, *Optimal experimental design*, WIREs Computational Statistics, 2 (2010), pp. 581–589.
- [8] PHILIP J. MAYBANK AND JONATHAN P. WHITELEY, *Automatic simplification of systems of reaction-diffusion equations by a posteriori analysis*, Mathematical Biosciences, 248 (2014), pp. 146–157.
- [9] LINDA PETZOLD, SHENGTAI LI, YANG CAO, AND RADU SERBAN, *Sensitivity analysis of differential-algebraic and partial differential equations*, Computers and Chemical Engineering, 30 (2006), pp. 1553–1559.
- [10] LUC PRONZATO, *Optimal experimental design and some related control problems*, Automatica, 44 (2008), pp. 303–325.
- [11] CHRISTIAN R. ROJAS, JAMES S. WELSH, GRAHAM C. GOODWIN, AND ARIE FEUER, *Robust optimal experiment design for system identification*, Automatica, 43 (2007), pp. 993–1008.
- [12] CLAY SHONKWILER, *Principal angles in terms of inner products*.