MATH 3795
Lecture 11. Applications of LLS. Data Assimilation.

Dmitriy Leykekhman

Goals

▶ Estimating the boundary data from observations.
Data Assimilation.

- We want to determine the initial conditions of a differential equation or system of differential equations from measurements of the solution at certain points and times.
- Such problems arise, for example, in weather forecasting.
- The weather (temperature, humidity, wind velocity,...) can be modeled using systems of partial differential equations.
- If we knew the weather (temperature, humidity, wind velocity,...) at every point in space, say, for Jan. 1, 12:00am, then we can solve the partial differential equations and use the computed solution to forecast the weather.
- Unfortunately, we can only measure temperature, humidity, wind velocity,... at a few locations. We need to use these measurements to estimate the weather (temperature, humidity, wind velocity,...) at every point in space, say, for Jan. 1, 12:00am (we estimate the initial condition).
Data Assimilation.

Consider simpler model problem so-called advection-diffusion equation

\[
\frac{\partial}{\partial t} y(x, t) = \nu \frac{\partial^2}{\partial x^2} y(x, t) - a \frac{\partial}{\partial x} y(x, t), \quad x \in (0, 1), \ t \in (0, T),
\]

\[
y(0, t) = 0, \quad t \in (0, T),
\]

\[
y(1, t) = 0, \quad t \in (0, T),
\]

\[
y(x, 0) = y_0(x), \quad x \in (0, 1).
\]

(1)

Here \( \nu \) is the diffusion parameter and \( a \) is the advection.

We want to determine the initial conditions \( y_0 \) from measurements of the solution \( y(x, t) \) at certain points and times.

We discretize the problem in space using the techniques that we have applied to solve BVPs in Assignment 3.
Data Assimilation.

Let $0 = x_0 < x_1 < \cdots < x_n = 1$, $x_i = ih$, and $h = \frac{1}{n+1}$.

Discretization in space gives us a system of ODE’s,

$$ y'(t) = Ky(t), \quad t \in (0, T], \quad y(0) = y_0, \quad (2) $$

where

$$ K = \begin{pmatrix}
-2\nu h^{-2} - ah^{-1} & \nu h^{-2} \\
\nu h^{-2} + ah^{-1} & -2\nu h^{-2} - ah^{-1} & \nu h^{-2} \\
& \ddots & \ddots & \ddots \\
& & & \ddots & \ddots & \ddots
\end{pmatrix} $$

and

$$ y_i(0) = y_0(x_i), \quad i = 0, 1, \ldots, n. $$
Data Assimilation.

From theory of linear system of ODEs with constant coefficients, given any initial data $y_0$, the solution of the ODE system (2) is given by

$$y(t) = \exp(Kt)y_0;$$

where $\exp(Kt) \in \mathbb{R}^{n \times n}$ is the matrix exponential of $Kt$.

It can be evaluated using Matlab’s build in function $\expm$, e.g., $\expm(K \ast t)$.

NOTE: $\exp(K \ast t)$ is different from $\expm(K \ast t)$ and the former does not give the matrix exponential, but evaluates the exponential of the matrix entries.
Data Assimilation.

Let $a = 1$, $\nu = 0.01$ and $n = 100$ ($h = 1/101$).

The solution of (2) with initial condition $y_0 = (\sin(\pi h), \ldots, \sin(n\pi h))^T$ (which is the approximate solution of (1), with initial condition $y_0(x) = \sin(\pi x)$ in (1)) is shown below.

Solution of $y'(t) = Ky(t);\ y(0) = \sin((1:n)\pi h)$
Data Assimilation.

- Now, suppose that we do not know $y_0$.
- We want to estimate $y_0$ from measurements of $k$ components $y_{n/k}(t), y_{2n/k}(t), \ldots$ of the solution $y(t)$ of (2) at times $t_1 = 0.02, t_2 = 0.04, \ldots, t_m = 0.5$. In the next numerical experiment we use $k = 5$, i.e., we observe 5 out of $n = 100$ components.
- That is, we want to determine $y_0$ from measurements $z_1, \ldots, z_m$ of $Hy(t_1), \ldots, Hy(t_m)$, where $y$ is the solution of (2) with the unknown initial condition and $H \in \mathbb{R}^{k \times n}$ is the matrix with entries
  \[ H_{ij} = 1 \quad \text{if} \quad j = (n/k)i; \quad H_{ij} = 0 \quad \text{else}. \]
- The solution $y(t)$ of (2) is given by
  \[ y(t) = \exp(Kt)y_0. \]
  Hence, $Hy(t_i) = H \exp(Kt_i)y_0$ for $i = 1, \ldots, m$.
- We arrive at the least squares problem
  \[
  \min_{y_0} \left\| \begin{pmatrix}
    H \exp(Kt_1) \\
    \vdots \\
    H \exp(Kt_m)
  \end{pmatrix}
  y_0 - \begin{pmatrix}
    z_1 \\
    \vdots \\
    z_m
  \end{pmatrix}
  \right\|_2^2.
  \]
Data Assimilation.

Thus we have a least square problem

\[
\min_{\mathbf{y}_0} \| A \mathbf{y}_0 - \mathbf{b} \|^2_2,
\]

where

\[
A = \begin{pmatrix}
H \exp(Kt_1) \\
\vdots \\
H \exp(Kt_m)
\end{pmatrix} \in \mathbb{R}^{km \times n}, \quad b = \begin{pmatrix}
z_1 \\
\vdots \\
z_m
\end{pmatrix} \in \mathbb{R}^{km}.
\]

Remember, at each time \( t_1 = 0.02, t_2 = 0.04, \ldots, t_m = 0.5 \) we observe \( k \) components of \( y(t) \). Consequently, we have \( km \) measurements

\[
H \exp(Kt_i) \in \mathbb{R}^{k \times n} \quad \text{and} \quad z_i \in \mathbb{R}^k \quad \text{for} \quad i = 1, \ldots, m.
\]
Data Assimilation.

Let \( a = 1, \nu = 0.01 \) and \( n = 100 \) (\( h = 1/101 \)). We solve the (2) with initial condition \( y_0^{ex} = (\sin(\pi h), \ldots, \sin(n\pi h))^T \) and at each time \( t_1 = 0.02, t_2 = 0.04, \ldots, t_m = 0.5 \). We record \( k = 5 \) components. This gives \( k \times m = 5 \times 25 = 125 \) measurements. The solution \( y_0 \) of the least squares problem is shown below.
Data Assimilation.

- Why is the computed solution so bad?
Data Assimilation.

- Why is the computed solution so bad?
- Recall that if $A = U \Sigma V^T$ is the SVD of $A \in \mathbb{R}^{km \times n}$, then the minimum norm solution of $\min \| Ay_0 - (b_{ex} + \delta b) \|_2^2$ is

$$y_0 = \sum_{i=1}^{r} \left( \frac{u_i^T b_{ex}}{\sigma_i} + \frac{u_i^T \delta b}{\sigma_i} \right) v_i,$$

If a singular value $\sigma_i$ is small, then $\frac{u_i^T \delta b}{\sigma_i}$ could be large, even if $u_i^T \delta b$ is small. This shows that errors $\delta b$ in the data can be magnified by small singular values $\sigma_i$. 
Data Assimilation.

- Why is the computed solution so bad?
- Recall that if $A = U \Sigma V^T$ is the SVD of $A \in \mathbb{R}^{k \times m \times n}$, then the minimum norm solution of $\min \|Ay_0 - (b_{ex} + \delta b)\|_2^2$ is

\[ y_0 = \sum_{i=1}^r \left( \frac{u_i^T b_{ex}}{\sigma_i} + \frac{u_i^T \delta b}{\sigma_i} \right) v_i, \]

If a singular value $\sigma_i$ is small, then $\frac{u_i^T \delta b}{\sigma_i}$ could be large, even if $u_i^T \delta b$ is small. This shows that errors $\delta b$ in the data can be magnified by small singular values $\sigma_i$.

- In our case, the errors in the data are due to floating point errors in computing

\[ Hy(t_i) = H\exp(Kt_i)y_0 \quad \text{for} \quad i = 1, \ldots, m. \]
Data Assimilation.

![Singular values of A graph](image)
To compute better estimates of the initial data $y_0$, we consider the regularized least squares problem

$$
\min_{y_0} \frac{1}{2} \| Ay_0 - (b + \delta b) \|_2^2 + \frac{\lambda}{2} \| y_0 \|_2^2
$$

for $\lambda \geq 0$. Let $y_{0,\lambda}$ denote the solution of this regularized LLS problem.

We use the data as before, but now add a 1% random error to the right hand side

$$
db = 0.01 \times \text{rand(size(b)))} \times b;
$$

$$
b = b + db;
$$

For various values of $\lambda$, we compute the error $\| y_{0,\lambda}^{ex} - y_{0,\lambda} \|_2$ as well as the residual $\| Ay_{0,\lambda} - b \|_2$. 
Data Assimilation.

The error $\| y_{ex} - y_{0,\lambda} \|_2$ for different values of $\lambda$ (log-log-scale)

The residual $\| Ay_0,\lambda - b \|_2$ and $\| \delta b \|_2$ for different values of $\lambda$ (log-log-scale)
Data Assimilation.

Using the bisection method (will study later) we compute an approximate solution \( \lambda \geq 0 \) of

\[
\| Ay_{0,\lambda} - (b + \delta b) \|_2^2 - \| \delta b \|_2^2 = 0.
\]

The computed root is \( \lambda \approx 0.0034 \). The solution \( y_{0,\lambda} \) is plotted below.