Space Mapping and Defect Correction

P.W. Hemker    D. Echeverría

Amsterdam

Lorentz Center    Leiden, sept 2005
Contents:

Space mapping (SM)
- space mapping idea
- 2 space mapping solutions
- perfect mapping
- reachable designs & flexible models

Defect correction (DeC)
- solution of equations \(\Rightarrow\) optimization problems
- SM is a special case
- left- and right- preconditioning
- improved space mapping

Examples
Two models:

Specifications of the aim: \( y \)

Fine model \( f(x) \) find design \( x^* : f(x^*) \approx y \)

Coarse model \( c(z) \) find design \( z^* : c(z^*) \approx y \)

\( f(x) \) gives a much more accurate result
\( c(z) \) is much simpler to solve

Use the coarse model to efficiently solve the fine model problem
Two models:

Specifications of the aim: \( y \)

Fine model
\[ f(x) \quad \text{find design } x^* : \quad f(x^*) \approx y \]

Coarse model
\[ c(z) \quad \text{find design } z^* : \quad c(z^*) \approx y \]

\( f(x) \) gives a much more accurate result
\( c(z) \) is much simpler to solve

\[
x^* = \arg \min_{\xi \in X} \| f(\xi) - y \|
\]

\[
z^* = \arg \min_{\zeta \in Z} \| c(\zeta) - y \|
\]
Reachable designs

\[ y \in f(X) : \text{y is a fine-model reachable design:} \]

\[ y \in c(Z) : \text{y is a coarse-model reachable design:} \]
**Reachable designs**

- **f-designs** $X$
- **c-designs** $Z$

Possible aims $f(X)$ and responses $c(Z)$ are connected by functions $f$ and $c$.

Mathematical notation:

- $X, Z \subset \mathbb{R}^n$: design
- $y \in \mathbb{R}^m$: aim
- $n$: number of degrees of freedom
- $m$: number of requirements

- $m > n$ / $m < n$: over- / under-determined systems

Generally $m > n$.
Reachable aim

\[
f(x^*) = y \\
c(z^*) = y
\]
Non-reachable aim

\[ x^* = \arg\min_{\xi} \| f(\xi) - y \| \]
\[ z^* = \arg\min_{\zeta} \| c(\zeta) - y \| \]
Misalignment function

\[ f(x) \approx c(z) \]

\[ r(z, x) = \| c(z) - f(x) \| \]
Space mapping function

\[ r(z, x) = \| c(z) - f(x) \| \]

\[ p(x) = \arg\min_{z \in Z} \| c(z) - f(x) \| \]

\( p(x) \): the best coarse-model control \( z \in Z \) yielding the response like \( x \in X \) for the fine-model.
Space mapping

\[ p(x) = \arg\min_{z \in Z} \| c(z) - f(x) \| \]

\[ z^* \approx p(x^*) \]

\[ y \approx c(p(x^*)) \]

\[ \Rightarrow \]

\[ c(p(x)) \approx f(x) \]

surrogate model
Original space mapping idea

\[ x^* \approx p^{-1}(z^*) \approx p^{-1}(c^{-1}(y)) \]
Inverse of $p(x)$

\[ p(x) = \arg\min_{z \in Z} \| c(z) - f(x) \| \]

\[ p : X \rightarrow Z, \quad p^{-1} : Z \rightarrow X \]

existence of $p^{-1} : Z \rightarrow X$ is not guaranteed !!!
Inverse of $p(x)$

\[ p(x) = \arg\min_{z \in Z} \| c(z) - f(x) \| \]

$p : X \to Z$, \quad $p^{-1} : Z \to X$

existence of $p^{-1} : Z \to X$ is not guaranteed !!!
Inverse of $p(x)$

difficulty for the original space mapping idea!
Two space mapping solutions

\[ x^* \approx p^{-1}(z^*) \approx p^{-1}(c^{-1}(y)) \]

\[ c(p(x)) \approx f(x) \quad \text{surrogate} \]

\[ p(x^*) \approx z^* \]
\[ x_p^* = \arg\min_x \| p(x) - z^* \| \]

\[ c(p(x^*)) \approx y \]
\[ x_d^* = \arg\min_x \| c(p(x)) - y \| \]
Perfect mapping

traditional space mapping theory:

The SM function $p$ is called a perfect mapping iff

$$z^* = p(x^*) ,$$

strange notion !!!

"perfect mapping" is NOT a property of $p$ alone !
Lemma 1

if \( y \) is \( f \)-reachable then any space mapping is perfect

\[ y \in f(X) \implies z^* = p(x^*) \]
Summary space mapping approaches

(0) original SM

\[ x_{sm}^* = p^{-1}(z^*) \]

(1) primal SM  (nearest approximation)

\[ x^*_p = \arg\min_x \|p(x) - z^*\|_Z \]

(2) dual SM  (best fitting to surrogate)

\[ x^*_d = \arg\min_x \|c(p(x)) - y\|_Y \]
**Injection & surjection**

**Injection**

\[ \exists \text{ left-inverse } q : \quad q \circ p = I_X \]

**Surjection**

\[ \exists \text{ right-inverse } q : \quad p \circ q = I_Y \]

\[ \exists \text{ inverse } q = p^{-1} \]

**Bijection** \( \equiv \) injective & surjective \( \equiv \) one-to-one
Lemma 2

\( z^* \in p(X) \Rightarrow p(x^*_p) = p(x^*_d) \)

**Additional assumptions needed:**

\( p \) injective and perfect \( \Rightarrow \) \( x^*_p = x^*_d = x^* \)
Flexibility of models

A coarse model is *more flexible* than a fine model:

\[ c(Z) \supset f(X) \]

A fine model is *more flexible* than a coarse model:

\[ f(X) \supset c(Z) \]

Models are *equally flexible*:

\[ f(X) = c(Z) \]
Lemma 3

\[ f(X) \supset c(Z) \quad \Rightarrow \quad \text{p surjective} \quad \Leftrightarrow \quad \exists \text{right-inverse} \quad pq = I_Z \]

\[ c(Z) \supset f(X) \quad \Rightarrow \quad \text{p injective} \quad \Leftrightarrow \quad \exists \text{left-inverse} \quad qp = I_X \]

equally flexible \quad \Rightarrow \quad p \text{ bijective} \quad \& \quad p \text{ perfect}
Defect Correction (classical)

hard to solve:

\[ \mathcal{F} x = y, \]

simplified version:

\[ \tilde{\mathcal{F}} x = y, \]

\[ x = \tilde{\mathcal{G}} y, \]

\[ \begin{cases} 
\quad x_0 = \tilde{\mathcal{G}} y, \\
\quad x_{k+1} = x_k - \tilde{\mathcal{G}} \mathcal{F} x_k + \tilde{\mathcal{G}} y,
\end{cases} \tag{A} \]

\[ \begin{cases} 
\quad \tilde{\mathcal{F}} x_0 = y, \\
\quad \tilde{\mathcal{F}} x_{k+1} = \tilde{\mathcal{F}} x_k - \mathcal{F} x_k + y.
\end{cases} \tag{B} \]
Defect Correction (classical)

hard to solve:
\[ \mathcal{F} x = y, \]
simplified version:
\[ \tilde{\mathcal{F}} x = y, \]
\[ x = \tilde{\mathcal{G}} y, \]

\[(A) \quad \begin{cases} 
  x_0 &= \tilde{\mathcal{G}} y, \\
  x_{k+1} &= x_k - \tilde{\mathcal{G}} \mathcal{F} x_k + \tilde{\mathcal{G}} y, 
\end{cases} \]

\[(B) \quad \begin{cases} 
  \tilde{\mathcal{F}} x_0 &= y, \\
  \tilde{\mathcal{F}} x_{k+1} &= \tilde{\mathcal{F}} x_k - \mathcal{F} x_k + y. 
\end{cases} \]
Defect Correction for optimization

Not surjective $\rightarrow$ left-inverse $G$

$G \mathcal{F} = I_X$

$\mathcal{F} x = y \iff f(x) = y$
$x = G y \iff x = \text{argmin} \|f(\xi) - y\|$

$\tilde{\mathcal{F}} x = y \iff c(\tilde{p}(x)) = y$
$x = \tilde{G} y \iff x = \text{argmin} \|c(\tilde{p}(\xi)) - y\|$

$\tilde{p}$ is **NOT** the space mapping function but a bijection arbitrary / easy to compute
Defect correction iteration

\[ z_0 = z^* = \arg\min_{\zeta \in \mathbb{Z}} \| c(\zeta) - y \| \]

\[ (A) \quad z_{k+1} = z_k - \arg\min_{\zeta \in \mathbb{Z}} \| c(\zeta) - f(\overline{q}(z_k)) \| + z^* \]

\[ (B) \quad z_{k+1} = \arg\min_{\zeta \in \mathbb{Z}} \| c(\zeta) - c(z_k) + f(\overline{q}(z_k)) - y \| \]

\[ x_{k+1} = \overline{q}(z_{k+1}) \]
Orthogonality relations

Convergence with limit \( \bar{x} = \lim_{k \to \infty} x_k, \quad \bar{z} = \lim_{k \to \infty} z_k \)

\[(A) \quad \rightarrow \quad f(\bar{x}) - y \in c(Z)^\perp(z^*) \]

\[(B) \quad \rightarrow \quad f(\bar{x}) - y \in c(Z)^\perp(\bar{z}) \]

but \(!!\) \( f(x^*) - y \in f(X)^\perp(x^*) \)

In other words, \(\bar{x}\) is found so that \(f(\bar{x})\) is closest to \(y\) seen from \(c(Z)\) !
Orthogonality relation 2D

\[ f(x^*) - y \in c(Z)^\perp(z^*) \]

\[ f(x^*) - y \in f(X)^\perp(x^*) \]
Orthogonality relation 3D

\[ f(x^*) - y \perp f(X)(x^*) \]
**Better Space Mapping (manifold mapping)**

\[ S : c(Z) \longrightarrow f(X) \]

\[ S(c(Z)) \cong f(X) \]

equally flexible!

surrogate model: \( S \circ c \circ \overline{p} \)

\[ S(c(\overline{p}(x))) \cong f(x) \]

\[ S v = S(c(\overline{z})) + S(v - c(\overline{z})) \]
Better space mapping (left and right)

\[ f \approx S_k \circ c \circ \overline{p} \]
Better space mapping

special case of defect correction

\[ X \longrightarrow Y \]

\[ F(x) = f(x) \quad \text{injection (no surjection)} \]
\[ G(y) = \operatorname{argmin}_\xi \| f(\xi) - y \| \quad \text{leftinverse} \]

\[ \tilde{F}_k(x) = S_k(c(p(x))) \]
\[ \tilde{G}_k(y) = \operatorname{argmin}_\xi \| S_k(c(p(\xi))) - y \| \]
Better space mapping algorithm

\[ x_0 = \arg\min_x \| c(\overline{p}(x)) - y \| \]

\[ k = 0, \]

\[ S_0 = I \]

for \( k = 0, 1, \cdots \)

do compute \( f(x_k) \) and \( c(\overline{p}(x_k)) \)

if \( k > 0 \) compute \( S_k \) such that for \( l = k - 1, k - 2, \cdots \),

\[ S_k (c(\overline{p}(x_l)) - c(\overline{p}(x_k))) = f(x_l) - f(x_k) \]

\[ x_{k+1} = \arg\min_{\xi} \| c(\overline{p}(\xi)) - c(\overline{p}(x_k)) - S_k^\dagger(y - f(x_k)) \| \]

enddo
**Conclusion**

- **Space mapping**: good principle to combine \( f(x) \) and \( c(z) \)
  - two types of ‘space mapping’ solution: \( x_p^* \neq x_d^* \neq x^* \)
  - ‘perfect mapping’ is a confusing notion
  - knowledge about **flexibility** is important
  - best is \( c(Z) = f(X) \)

- **Defect correction**: general framework
  - leads to (generalization of the) SM approach
  - suggests analysis / new variants

- **Better space mapping by left-preconditioning**
  - yields the true solution \( x^* \) !!!
  - faster solution for optimization problems
Example I (first data fitting problem; \( n = 2, m = 3 \))

\[
\begin{align*}
\mathbf{f}(\mathbf{x}) &= [(x_1 t_i)^2 + x_1 t_i + x_2]_{i=1,2,3} & t &= [-1, 0, 1] \\
\mathbf{c}(\mathbf{z}) &= [z_1 t_i + z_2]_{i=1,2,3} & \mathbf{y} &= [y_1, y_0, y_1] \\
\mathbf{p}(\mathbf{x}) &= [x_1, x_2 + \frac{2}{3} x_1^2]
\end{align*}
\]

bijection \( \Rightarrow \mathbf{x}_p = \mathbf{x}_d \)

perfect mapping: \( \mathbf{x}^* = \mathbf{x}_p^* = \mathbf{x}_d^* \) \( \Leftarrow \) coincide!

generally \( \mathbf{x}^* \neq \mathbf{x}_p^* = \mathbf{x}_d^* \)

Generally, the traditional SM-solution is not the accurate solution \( \mathbf{x}^* \)!
Example I (deviation from perfect mapping)

\[ y = f(\hat{x}) + n \quad (n \text{ perturbarion}) \]

Quality of the space mapping optimum depends on the deviation from perfect mapping.

\[ \| f(x^*) - y \|_2 \quad \| f(x_{sm}^*) - y \|_2 \]

\[ \| f(\hat{x}) \| = 0.23 \]
Example II (second data fitting problem; \( n = 2, m = 3 \))

\[
\begin{align*}
  f(x) & = [(x_1(x_2t_i + 1)^2]_{i=1,2,3} & t & = [-1, 0, 1] \\
  c(z) & = [z_1t_i + z_2]_{i=1,2,3} & y & = [y_{-1}, y_0, y_1] \\
  p(x) & = [x_1x_2, x_1(a + \frac{2}{3}x_2^2)] & \iff \text{NO bijection}
\end{align*}
\]

4 different possibilities, depending on \( y \)
Example II (second data fitting problem; \(n = 2, m = 3\))

\[
f(x) = [(x_1(x_2t_i + 1)^2)]_{i=1,2,3} \quad \quad \quad t = [-1, 0, 1]
\]

\[
c(z) = [z_1t_i + z_2]_{i=1,2,3} \quad \quad \quad y = [y_1, y_0, y_1]
\]

Four different data sets \(y\)

<table>
<thead>
<tr>
<th>Case</th>
<th>Condition</th>
<th>(z^* = p(x^*))</th>
<th>(x^* = x_p = x_d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(y = f(x))</td>
<td>reachable design</td>
<td>perfect mapping</td>
</tr>
<tr>
<td>2.</td>
<td>(y \neq f(x))</td>
<td>non-reachable design</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>(z \in p(X))</td>
<td>(z^* \neq p(x^*))</td>
<td>no perfect mapping</td>
</tr>
<tr>
<td>4.</td>
<td>(z \not\in p(X))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Traditional space mapping: depending on the data various situations occur. Better space mapping finds \(x^*\) in all cases!
**Example II, location of** \( z^* \), \( p(x^*) \), \( p(x^*_p) \), \( p(x^*_d) \)

Four data sets \( y \):
(1) and (2): coinciding \( z^* = p(x^*_p) = p(x^*_d) = p(x^*) \)
(3): coinciding \( p(x^*_p) = p(x^*_d) \) separate \( z \) and \( p(x^*) \)
(4): separate \( z^*, p(x^*_p), p(x^*_d) \) and \( p(x^*) \)
## Example II - results

<table>
<thead>
<tr>
<th></th>
<th>$[y_{-1}, y_0, y_1]$</th>
<th>$x^*$</th>
<th>$z^*$</th>
<th>$C(z^*)$</th>
<th>$F(z^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[0.081, 0.100, 0.121]</td>
<td>[0.100, 0.100]</td>
<td>[0.020, 0.101]</td>
<td>8.16e-04</td>
<td>1.41e-01</td>
</tr>
<tr>
<td>2</td>
<td>[0.100, 0.101, 0.102]</td>
<td>[0.101, 0.006]</td>
<td>[0.001, 0.101]</td>
<td>5.67e-06</td>
<td>1.73e-01</td>
</tr>
<tr>
<td>3</td>
<td>[0.000, -0.400, 0.100]</td>
<td>[-0.101, -0.141]</td>
<td>[0.050, -0.100]</td>
<td>3.67e-01</td>
<td>4.58e-01</td>
</tr>
<tr>
<td>4</td>
<td>[0.000, -0.350, 0.200]</td>
<td>[-0.059, -0.352]</td>
<td>[0.100, -0.050]</td>
<td>3.67e-01</td>
<td>4.76e-01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$x_p^*$</th>
<th>$x_d^*$</th>
<th>$F(x^*)$</th>
<th>$F(x_p^*)$</th>
<th>$F(x_d^*)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM</td>
<td>1</td>
<td>[0.100, 0.100]</td>
<td>[0.100, 0.100]</td>
<td>1.96e-17</td>
<td>3.10e-17</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>[0.101, 0.006]</td>
<td>[0.101, 0.006]</td>
<td>3.00e-06</td>
<td>3.00e-06</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>[-0.096, -0.261]</td>
<td>[-0.096, -0.261]</td>
<td>3.70e-01</td>
<td>3.73e-01</td>
</tr>
<tr>
<td></td>
<td>4 a</td>
<td>[-0.034, -1.225]</td>
<td>[-0.033, -1.225]</td>
<td>3.83e-01</td>
<td>4.12e-01</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>[0.007, 4.007]</td>
<td></td>
<td>3.63e-01</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\bar{x}$</th>
<th>$F(\bar{x})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[0.100, 0.100]</td>
<td>6.30e-06</td>
</tr>
<tr>
<td>2</td>
<td>[0.101, 0.005]</td>
<td>2.04e-06</td>
</tr>
<tr>
<td>3</td>
<td>[-0.101, -0.141]</td>
<td>3.70e-01</td>
</tr>
<tr>
<td>4</td>
<td>[-0.060, -0.304]</td>
<td>3.83e-01</td>
</tr>
</tbody>
</table>
Example II - Case (4)

- Global minimum of $F(x)$
- Minimum of $C(x)$
- Local minimum of $F(x)$
- $\bar{x}$
**Example III - magnetostatics**

A two-dimensional magnetostatics problem. \( n = 3, m = 3 \)

\( f(x) \) : finite element simulation
  (magnetostatic Maxwell / potential equation)
\( c(z) \) : magnetic circuit

\[ X \equiv Z \]

Magnetic flux specified at \( \bullet \)
**Example III - results**

<table>
<thead>
<tr>
<th></th>
<th># Evaluations</th>
<th>Final design (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(z^*)</td>
<td></td>
<td>[5.3571, 7.5000, 5.0000]</td>
</tr>
<tr>
<td>ASM</td>
<td>4</td>
<td>[7.9897, 7.5821, 6.5396]</td>
</tr>
<tr>
<td>SMI</td>
<td>6</td>
<td>[7.9797, 7.5808, 6.5381]</td>
</tr>
<tr>
<td>qN</td>
<td>6</td>
<td>[7.9891, 7.5821, 6.5394]</td>
</tr>
<tr>
<td>NMS</td>
<td>62</td>
<td>[8.0000, 7.5872, 6.5372]</td>
</tr>
<tr>
<td>DIRECT</td>
<td>186</td>
<td>[7.9806, 7.5823, 6.5432]</td>
</tr>
</tbody>
</table>

Efficiency comparison:
- ASM : the ASM algorithm;
- SMI : the simplified ASM algorithm with \(B_k = I\);
- qN : a quasi-Newton scheme;
- NMS : Nelder-Mead simplex
- DIRECT: a direct search method

Tolerance: \( \| p(x_k) - z^* \| < 0.001 \| z^* \| \)
**Example III - results**

The convergence history for the cost function in ASM iteration.

Design with the **nonlinear fine model** and different coarse models.

![Graph showing convergence history with different models](image)

**better coarse model ⇒ faster convergence**
Conclusion

• Space mapping: good principle to combine $f(x)$ and $c(z)$
  – two types of ‘space mapping’ solution: $x_p^* \neq x_d^* \neq x^*$
  – ‘perfect mapping’ is a confusing notion
  – knowledge about flexibility is important
  – best is $c(Z) = f(X)$

• Defect correction: general framework
  – leads to (generalization of the) SM approach
  – suggests analysis / new variants

• Better space mapping by left-preconditioning
  – yields the true solution $x^*$ !!!
  – faster solution for optimization problems
Primal SM Algorithm

\[ x_0 = z^* = \arg\min_z \|c(z) - y\| \]
\[ p_0 = I \]

for \( k = 0, 1, \ldots \)
do

\[ z_k = p(x_k) = \arg\min_z \|c(z) - f(x_k)\|_Y \]
\[ \{x_k, z_k\} \Rightarrow p_{k+1} \text{ approximation} \]

\[ x_{k+1} = \arg\min_x \|p_{k+1}(x) - z^*\| \]
endo
Dual SM Algorithm

\[ x_0 = z^* = \operatorname{argmin}_z \| c(z) - y \| \]
\[ p_0 = I \]

for \( k = 0, 1, \cdots \) do
\[ z_k = p(x_k) = \operatorname{argmin}_z \| c(z) - f(x_k) \|_Y \]
\[ \{ x_k, z_k \} \Rightarrow p_{k+1} \text{ approximation} \]
\[ x_{k+1} = \operatorname{argmin}_x \| c(p_{k+1}(x)) - y \| \]
endo
Aggressive SM Algorithm

\[ p_k(x) = p(x_k) + B_k(x - x_k) \]
**Aggressive SM Algorithm**

\[ p_k(x) = p(x_k) + B_k(x - x_k) \]

\[ B_{k+1} = B_k + \frac{p(x_{k+1}) - p(x_k) - B_k h^T}{h^T h}, \quad h = x_{k+1} - x_k \]

Broyden’s rank-one update
Aggressive SM Algorithm

\[ p_k(x) = p(x_k) + B_k(x - x_k) \]

\[ x_0 = z^* = \arg\min_z \|c(z) - y\| \]
\[ B_0 = I \]

for \( k = 0, 1, \ldots \)
do \[ z_k = p(x_k) = \arg\min_z \|c(z) - f(x_k)\| \]
\[ h = -B_k^\dagger(p(x_k) - z^*) \]
\[ x_{k+1} = x_k + h \]
\[ B_{k+1} = B_k + \frac{p(x_{k+1}) - p(x_k) - B_k h^T}{h^T h} \]
endo