Supplementary material for BSM 2024 poster Theory

Adrian Marin Mag[1], Paula Koelemeijer[1], and Christophe Zaroli[2]

[1]Department of Earth Sciences, University of Oxford, Oxford, UK
[2]Institut Terre et Environnement de Strasbourg, Université de Strasbourg, EOST, CNRS, UMR 7063, 5 rue Descartes, Strasbourg F-67084, France

1 Spaces

The model is composed of one or more physical parameters of interest (for example only v_p, or v_p, v_s, ρ). Each physical parameter lives in a space $\mathcal{M}_j$, which may be a $\mathbb{R}^N$ or some function space. For this poster we use as example physical parameters that are piece-wise continuous and bounded functions defined on some spatial domains Ω_j (in short notation $PCb[\Omega_j]$). We assume that the spaces are Hilbert spaces with some inner product defined on them. For example, in this poster we assume the following inner product:

$$\langle f, g \rangle_{\mathcal{M}_j} = \int_{\Omega_j} f g d\Omega_j \tag{1}$$

where $f, g \in \mathcal{M}_j$. The norm induced by the inner product is:

$$\|f\|_{\mathcal{M}_j} = \sqrt{\langle f, f \rangle_{\mathcal{M}_j}}. \tag{2}$$

When we have multiple physical parameters, each member of some $\mathcal{M}_j$, we construct the total model space $\mathcal{M}$ (also called simply the model space) by a direct sum:

$$\mathcal{M} = \bigoplus_j \mathcal{M}_j \tag{3}$$

Every member of $\mathcal{M}$ will therefore be a tuple $m = (m^1, m^2, ...)$. Such a construction does not impose any assumptions on the links between the various physical parameters m^j. The inner product and norm on $\mathcal{M}$ are defined by:

$$\langle m_1, m_2 \rangle_{\mathcal{M}} = \sum_j \left\langle m_1^j, m_2^j \right\rangle_{\mathcal{M}_j} \tag{4}$$

$$\|m\|_{\mathcal{M}} = \sum_j \left\| m^j \right\|_{\mathcal{M}_j} \tag{5}$$

where $m^1, m^2 \in \mathcal{M}$.

The data space $\mathcal{D}$ and the property space $\mathcal{P}$ will always be some $\mathbb{R}^N$, where N is either the number of data, or the number of property values. We assume the usual inner product and induced norm on $\mathbb{R}^N$:

$$\langle a, b \rangle_{\mathbb{R}^N} = \sum_i a_i b_i \tag{6}$$

$$\|a\|_{\mathbb{R}^N} = \sqrt{\sum_i a_i^2} \tag{7}$$

2 Mappings

There are two important mappings: the model-data mapping G and the property mapping $\mathcal{T}$. They have the general form:

$$d_i = G(m) = \sum_j \left\langle K_i^j, m^j \right\rangle_{\mathcal{M}_j} \tag{8}$$

$$p^{(k)} = \mathcal{T}(m) = \sum_j \left\langle T^{j,(k)}, m^j \right\rangle_{\mathcal{M}_j} \tag{9}$$

where K_i^j are the sensitivity kernels, and $T^{j,(k)}$ are the target kernels. The kernels are members of the spaces $\mathcal{M}_j$ and are assumed to be linearly independent.

3 Norm Bound

The prior information used in this method is, following (Al-Attar 2021):

$$\|m\|_{\mathcal{M}} \leq M \tag{10}$$

Usually the norm bound M is found by assuming that the absolute value of the true model $\bar{m}$ is bounded above by some functions b^j such that:

$$\left| m^j \right| \leq b_j \tag{11}$$

$$\|m\| = \sum_j \left\| m^j \right\| \leq \sum_j \|b_j\| = M \tag{12}$$

4 Solution

This overview of the solution derivation follows Al-Attar 2021. The problem that we want to solve is:

$$\begin{gathered}\text{Given}\\ G(m)=d \quad (13)\\ \|m\|_{\mathcal{M}} \le M \quad (14)\\ \text{Find}\\ \mathcal{T}(\bar{m})=\bar{p}. \quad (15)\end{gathered}$$

If G is surjective, but not injective, then there might be multiple models that fit the data perfectly, making an inversion ill-posed. A general solution of (13) can be given in terms of the generalized Moore-Penrose right inverse and the null space of G (also known as the kernel and denoted by *ker*):

$$\{m\}=\tilde{m}+\ker(G) \tag{16}$$

where $\tilde{m}$ is the generalized inverse (in this case it is also equal to the least norm solution of eq. (13)). Therefore we can write the true model solution as:

$$\bar{m}=\tilde{m}+m_0 \tag{17}$$

where $m_0 \in \ker(G)$. If we apply the property mapping we obtain:

$$\bar{p}=\tilde{p}+\mathcal{T}(m_0) \tag{18}$$

where $\tilde{p}$ is the property of the least norm solution $\tilde{m}$, and $\bar{p}$ is the property of the true model. While $\tilde{p}$ is known and fixed, $\mathcal{T}(m_0)$ may take many values. To find the range of values that $\mathcal{T}(m_0)$ we will do the following operations: replace (17) into (14) squared:

$$\|\bar{m}\|^2_{\mathcal{M}}=\langle\bar{m},\bar{m}\rangle_{\mathcal{M}}=\langle\tilde{m}+m_0,\tilde{m}+m_0\rangle\le M^2 \tag{19}$$

then use the fact that $\tilde{m}\perp m_0$ (see in next section why):

$$\|\tilde{m}\|^2_{\mathcal{M}}+\langle m_0,m_0\rangle_{\mathcal{M}}\le M^2 \tag{20}$$

Now, $m_0 \in \ker(G)$, and we note that a part of m_0 may be in $\ker(\mathcal{T})$. We call that part m_{00} and we can obtain it via:

$$m_{00}=\mathbb{P}_{\ker(\mathcal{T}_{|\ker(G)})}m_0 \tag{21}$$

where $\mathcal{T}_{|\ker(G)}$ is the restriction of $\mathcal{T}$ onto $\ker(G)$, and $\mathbb{P}_{\ker(\mathcal{T}_{|\ker(G)})}$ is the projection onto $\mathcal{T}_{|\ker(G)}$ operator (see next section to more details). We then separate m_0 into its two components in (20):

$$\|\tilde{m}\|^2_{\mathcal{M}}+\langle m_0-m_{00},m_0-m_{00}\rangle_{\mathcal{M}}+\|m_{00}\|^2_{\mathcal{M}}\le M^2 \tag{22}$$

where we used the fact that $m_0-m_{00}\perp m_{00}$ (see next section why). Replacing (21) into (22):

$$\begin{aligned}\Big\langle(1-\mathbb{P}_{\ker(\mathcal{T}_{|\ker(G)})})m_0,(1-\mathbb{P}_{\ker(\mathcal{T}_{|\ker(G)})})m_0\Big\rangle_{\mathcal{M}}\le\\ M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}}.\end{aligned} \tag{23}$$

However:

$$\begin{aligned}1-\mathbb{P}_{\ker(\mathcal{T}_{|\ker(G)})}=\\ 1-(1-\mathcal{T}^*_{|\ker(G)}(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}\mathcal{T}_{|\ker(G)})=\end{aligned} \tag{24}$$

$$\mathcal{T}^*_{|\ker(G)}(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}\mathcal{T}_{|\ker(G)} \tag{25}$$

(see next section for more details). Plugging this back into (23), and writing the inner product as a norm:

$$\begin{aligned}\left\|\mathcal{T}^*_{|\ker(G)}(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}\mathcal{T}_{|\ker(G)}(m_0)\right\|^2_{\mathcal{M}}\le\\ M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}}.\end{aligned} \tag{26}$$

However:

$$\mathcal{T}_{|\ker(G)}(m_0)=\mathcal{T}(m_0)=\mathcal{T}(\bar{m}-\tilde{m})=\bar{p}-\tilde{p} \tag{27}$$

which means that (26) can also be written as:

$$\begin{aligned}\left\|\mathcal{T}^*_{|\ker(G)}(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}(\bar{p}-\tilde{p})\right\|^2_{\mathcal{M}}\le\\ M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}}.\end{aligned} \tag{28}$$

Now let:

$$\mathcal{T}^*_{|\ker(G)}(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}=A \tag{29}$$

then (28) is:

$$\|A(\bar{p}-\tilde{p})\|^2_{\mathcal{M}}\le M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}} \tag{30}$$

$$\langle A(\bar{p}-\tilde{p}),A(\bar{p}-\tilde{p})\rangle^2_{\mathcal{M}}\le M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}} \tag{31}$$

$$\langle A^*A(\bar{p}-\tilde{p}),\bar{p}-\tilde{p}\rangle^2_{\mathcal{M}}\le M^2-\|\tilde{m}\|^2_{\mathcal{M}}-\|m_{00}\|^2_{\mathcal{M}}. \tag{32}$$

Using (29) we have:

$$A^*A=(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1} \tag{33}$$

which finally yields the result:

$$\Big\langle(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}(\bar{p}-\tilde{p}),\bar{p}-\tilde{p}\Big\rangle^2_{\mathcal{M}}\le M^2-\|\tilde{m}\|^2_{\mathcal{M}} \tag{34}$$

where $\|m_{00}\|^2_{\mathcal{M}}$ was omitted since it cannot be computed in practice (and it immediately shows that this innequality is not sharp).

Inequality (34) tells us that the true property $\bar{p}$ is found in a hyperellipsoid centered on $\tilde{p}$ and with a shape determined by $\mathcal{H}=(\mathcal{T}_{|\ker(G)}\mathcal{T}^*_{|\ker(G)})^{-1}$.

References

Al-Attar, David (2021). "Linear inference problems with deterministic constraints". In: *arXiv preprint arXiv:2104.12256*.