Spatial Varying Coefficient Model

Yewen Chen

2019/6/22

Contents

1 Hierarchical modeling for univariate spatial data

1.1 Model framework

Consider a simple linear model:

$$Y(s) = \mu(s) + \underbrace{w(s) + \varepsilon(s)}_{\eta(s)} \tag{1.1.1}$$

where the mean structure $\mu(s) = X(s)\boldsymbol{\beta}$, the residual $\eta(s)$ is partitioned into two pieces, one spatial and one nonspatial.

Assumptions of model (1.1.1):

a. $\varepsilon(s) \overset{i.i.d}{\sim} N(0, \tau^2)$;

b. $\boldsymbol{w}$ is a stationary process independent of $\boldsymbol{\varepsilon}$, and is a Gaussian processes(GP): $\boldsymbol{w} \sim GP(\mathbf{0}, \sigma^2 \boldsymbol{H}(\phi))$, where

$$(\boldsymbol{H}(\phi))_{ij} = \sigma^2 \rho(\phi; \|s_i - s_j\|) \quad i, j = 1, 2, \cdots, n.$$

1.2 Model decomposition

There are two ways to fit model from Bayes's point of view:

(a) $p(\boldsymbol{\theta}, \boldsymbol{w}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{w}, \boldsymbol{\beta}, \tau^2) p(\boldsymbol{w}|\sigma^2, \phi) \pi(\boldsymbol{\theta})$

(b) $p(\boldsymbol{\theta}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta})$

The structure of the model is so complex that **numerical algorithm** is needed.

2 Model Solution

2.1 MCMC algorithm

Turn to (a), our goal is to sample $\boldsymbol{\theta}, \boldsymbol{w}$ from $p(\boldsymbol{\theta}, \boldsymbol{w}|\boldsymbol{y})$ based on Gibbs sampler. Therefore, we need to solve $p\left(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{w}, \tau^2\right), p(\boldsymbol{w}|\boldsymbol{y}, \boldsymbol{\theta}), p\left(\tau^2|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\beta}, \boldsymbol{w}\right), p\left(\sigma^2|\boldsymbol{\phi}, \boldsymbol{w}\right)$ and $p\left(\phi|\boldsymbol{w}, \sigma^2\right)$.

Step1: Set $\pi(\boldsymbol{\beta})=N\left(\boldsymbol{A}\boldsymbol{\alpha},\boldsymbol{\Sigma}_{\beta}\right)$, then $p\left(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{w},\tau^2\right)=N(\boldsymbol{D}\boldsymbol{\eta},\boldsymbol{D})$, where

$$\boldsymbol{D}=\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{\tau^2}+\boldsymbol{\Sigma}_{\beta}^{-1}\right)^{-1};\boldsymbol{\eta}=\frac{\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{w})}{\tau^2}+\boldsymbol{\Sigma}_{\beta}^{-1}\boldsymbol{A}\boldsymbol{\alpha} \tag{2.1.1}$$

Step2: Since $\pi(\boldsymbol{w})=N\left(0,\sigma^2\boldsymbol{H}(\phi)\right)$, then $p(\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta})$ is again of the form $N(\boldsymbol{D}\boldsymbol{\eta},\boldsymbol{D})$, where

$$\boldsymbol{D}=\left(\frac{\boldsymbol{I}}{\tau^2}+\frac{\boldsymbol{H}^{-1}(\phi)}{\sigma^2}\right)^{-1};\boldsymbol{\eta}=\frac{(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})}{\tau^2} \tag{2.1.2}$$

Step3: Furthermore, set $\pi\left(\tau^2\right)=IG\left(a_\tau,b_\tau\right),\pi\left(\sigma^2\right)=IG\left(a_\sigma,b_\sigma\right)$, respectively, then

$$p\left(\tau^2|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta},\boldsymbol{w}\right)=IG\left(a_\tau+\frac{n}{2},b_\tau+\frac{(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})}{2}\right) \tag{2.1.3}$$

$$p\left(\sigma^2|\phi,\boldsymbol{w}\right)=IG\left(a_\sigma+\frac{n}{2},b_\sigma+\frac{\boldsymbol{w}'\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}}{2}\right) \tag{2.1.4}$$

Step4: However, no closed form is available for $p\left(\phi|\boldsymbol{w},\sigma^2\right)$ as follow

$$p\left(\phi|\boldsymbol{w},\sigma^2\right)\propto\pi(\phi)\exp\left(-\frac{\boldsymbol{w}'\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}}{2\sigma^2}\right) \tag{2.1.5}$$

Except ϕ' s sample, the other can be sampled by Gibbs sampler, while the former can be taken by Metropolis algorithm or slice sampling.

2.2 Variational Bayes

Give initial values to the expectation of $1/\tau^2,\phi,\mathbf{w}$ and $\mathbf{R}(\phi)^{-1}$: $\mathbf{E}^{(0)}\left(1/\tau^2\right)=\left(1/\tau^2\right)^{(0)},\mathbf{E}^{(0)}(\phi)=\phi^{(0)},\boldsymbol{\mu}_w^{(0)}=\mathbf{0}$ and $\mathbf{E}^{(0)}\left(\mathbf{R}(\phi)^{-1}\right)=\mathbf{R}\left(\phi^{(0)}\right)^{-1}$.

Step1: Update the distribution of $\boldsymbol{\beta}\sim\mathrm{MVN}\left(\mu_\beta^{(t)},\mathbf{V}_\beta^{(t)}\right)$, where

$$\mathbf{V}_\beta^{(t)}=\left[\boldsymbol{E}^{(t-1)}\left(1/\tau^2\right)\left(\mathbf{X}'\mathbf{X}\right)+\Sigma_\beta^{-1}\right]^{-1}$$

and

$$\boldsymbol{\mu}_\beta^{(t)}=\left[\boldsymbol{E}^{(t-1)}\left(1/\tau^2\right)\left(\mathbf{X}'\mathbf{X}\right)+\Sigma_\beta^{-1}\right]^{-1}\left[\boldsymbol{E}^{(t-1)}\left(1/\tau^2\right)\mathbf{X}'\left(\mathbf{Y}-\boldsymbol{\mu}_{\mathbf{w}}^{(t-1)}\right)+\Sigma_\beta^{-1}\right];$$

Step2: Update the distribution of $\tau^2 \sim IG$ with parameters $a_\tau + \frac{n}{2}$ and

$$b_\tau + \frac{1}{2}\left[\mathbf{tr}\left(\mathbf{V}_{\mathbf{w}}^{(t-1)}\right) + p\mathrm{E}^{(t-1)}\left(1/\tau^2\right) + \left(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{w}}^{(t-1)}\right)'\left(\mathbf{I}_n - \mathbf{H}\right)\left(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{w}}^{(t-1)}\right)\right],$$

where $\mathbf{H} = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'$, calculate $m_{\tau^2}^{(t)} = \mathrm{E}^{(t)}\left(1/\tau^2\right);$

Step3: Update the distribution of $\sigma^2 \sim IG$ with parameters $a_\sigma + \frac{n}{2}$ and

$$b_\sigma + \frac{1}{2}\left\{\mathbf{tr}\left[\mathrm{E}^{(t-1)}\left(\mathbf{R}(\phi)^{-1}\right)\mathbf{V}_{\mathbf{w}}^{(t-1)}\right] + \boldsymbol{\mu}_{\mathbf{w}}^{(t-1)'}\mathrm{E}^{(t-1)}\left(\mathbf{R}(\phi)^{-1}\right)\boldsymbol{\mu}_{\mathbf{w}}^{(t-1)}\right\};$$

calculate $m_{\sigma^2}^{(t)} = \mathrm{E}^{(t)}\left(1/\sigma^2\right);$

Step4: Update the distribution of $\mathbf{w} \sim \mathrm{MVN}\left(\mu_{\mathbf{w}}^{(t)}, \mathbf{V}_{\mathbf{w}}^{(t)}\right)$, where

$$\mathbf{V}_{\mathbf{w}}^{(t)} = \left[m_{\sigma^2}^{(t)}\mathrm{E}^{(t-1)}\left(\mathbf{R}(\phi)^{-1}\right) + m_{\tau^2}^{(t)}\mathbf{I}_n\right]^{-1}$$

and

$$\boldsymbol{\mu}_{\mathbf{w}}^{(t)} = m_{\tau^2}^{(t)}\left[m_{\sigma^2}^{(t)}\mathrm{E}^{(t-1)}\left(\mathbf{R}(\phi)^{-1}\right) + m_{\tau^2}^{(t)}\mathbf{I}_n\right]^{-1}\left(\mathbf{Y} - \mathbf{X}\boldsymbol{\mu}_\beta^{(t)}\right)$$

Step5: Update the distribution of ϕ which is proportional to

$$g(\phi) = |\mathbf{R}(\phi)|^{-\frac{1}{2}}\exp\left\{-\frac{m_{\sigma^2}^{(t)}\left[\mathrm{tr}\left(\mathbf{R}(\phi)^{-1}\mathbf{V}_{\mathrm{w}}^{(t)}\right) + \boldsymbol{\mu}_{\mathrm{w}}^{(t)}\mathbf{R}(\phi)^{-1}\boldsymbol{\mu}_{\mathrm{w}}^{(t)}\right]}{2}\right\} \tag{2.2.1}$$

and calculate $\mathrm{E}^{(t)}(\phi)$ and $\mathrm{E}^{(t)}\left(\mathbf{R}(\phi)^{-1}\right)$. However the distribution function (2.2.1) is not analytically tractable, so importance sampling is proposed to approximate.

$$\mathrm{E}(f(\phi)) = \frac{\int f(\phi)g(\phi)\mathrm{d}\phi}{\int g(\phi)\mathrm{d}\phi} = \frac{\int f(\phi)\frac{g(\phi)}{p_I(\phi)}p_I(\phi)\mathrm{d}\phi}{\int \frac{g(\phi)}{p_I(\phi)}p_I(\phi)\mathrm{d}\phi} \approx \frac{\frac{1}{N}\sum_{i=1}^N f\left(\phi_i\right)W\left(\phi_i\right)}{\frac{1}{N}\sum_{i=1}^N W\left(\phi_i\right)} = \sum_{i=1}^N f\left(\phi_i\right)W^*\left(\phi_i\right) \tag{2.2.2}$$

where $\phi_i \overset{\text{iid}}{\sim} p_I(\phi), W\left(\phi_i\right) = g\left(\phi_i\right)/p_I\left(\phi_i\right)$ and $W^*\left(\phi_i\right) = \frac{W(\phi_i)}{\sum_{i=1}^N W(\phi_i)}$.

(1) Because the distributions of the parameters other than ϕ only depend on $\mathrm{E}\left(\mathbf{R}(\phi)^{-1}\right)$ using the expectation from importance sampling in (2.2.2) allows the VB algorithm to proceed toward convergence.

(2) After the VB algorithm converges, importance sampling resampling method (Rubin, 1987) is used to simulate samples of ϕ, which is proportional to (2.2.1). Inferences about $p(\phi|\mathbf{Y})$ can be made based on these samples.

2.3 EM algorithm

E-step:

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) &= E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}\left(\log p(\boldsymbol{y},\boldsymbol{w}|\boldsymbol{\theta})\right) \\ &= E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}\left(\log p(\boldsymbol{y}|\boldsymbol{w},\boldsymbol{\theta})p(\boldsymbol{w}|\sigma^2,\phi)\right) \\ &= \underbrace{E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}\left(\log p(\boldsymbol{y}|\boldsymbol{w},\boldsymbol{\theta})\right)}_{Q_1} + \underbrace{E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}\left(\log p(\boldsymbol{w}|\sigma^2,\phi)\right)}_{Q_2} \end{aligned}$$

where $p(\boldsymbol{y}|\boldsymbol{w},\boldsymbol{\theta}) \sim N(\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w},\tau^2\boldsymbol{I})$; $p(\boldsymbol{w}|\sigma^2,\phi) \sim N(\boldsymbol{0},\sigma^2\boldsymbol{H}(\phi))$. Let $E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}(\cdot) = E(\cdot)$, then

$$Q_1 = -\frac{n}{2}\log\tau^2 - \frac{E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})}{2\tau^2}$$

$$Q_2 = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\log|\boldsymbol{H}(\phi)| - \frac{E[\boldsymbol{w}'\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}]}{2\sigma^2}.$$

M-step: Therefore, to maximize Q_1 and Q_2 and then obtain

$$\hat{\tau}^2 = \frac{(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}) - 2(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'E(\boldsymbol{w}) + E(\boldsymbol{w}'\boldsymbol{w})}{n} \tag{2.3.1}$$

$$\hat{\sigma}^2 = \frac{E[\boldsymbol{w}'\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}]}{n} \tag{2.3.2}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y}-E(\boldsymbol{w})) \tag{2.3.3}$$

$$\phi_{k+1} = \phi_k - \frac{\partial\boldsymbol{Q}_2/\partial\phi|_{\phi=\phi_k}}{\partial^2\boldsymbol{Q}_2/\partial\phi^2|_{\phi=\phi_k}} \tag{2.3.4}$$

where

$$\begin{aligned} E\left[\boldsymbol{w}'\boldsymbol{w}\right] &= \mathbf{tr}(\boldsymbol{D}) + (\boldsymbol{D}\boldsymbol{\eta})'\boldsymbol{D}\boldsymbol{\eta}; \\ E\left[\boldsymbol{w}'\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}\right] &= \mathbf{tr}\left(\boldsymbol{H}^{-1}(\phi)E\left[\boldsymbol{w}\boldsymbol{w}'\right]\right) \\ &= \mathbf{tr}\left(\boldsymbol{H}^{-1}(\phi)(\boldsymbol{D}+\boldsymbol{D}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{D}')\right), \end{aligned}$$

see (2.1.2) for $\boldsymbol{D}$ and $\boldsymbol{\eta}$. In addition,

$$\frac{\partial\boldsymbol{Q}_2}{\partial\phi} = -\frac{1}{2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial\boldsymbol{H}}{\partial\phi}\right) - \frac{1}{2\sigma^2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial\boldsymbol{H}}{\partial\phi}\boldsymbol{H}^{-1}(\boldsymbol{D}+\boldsymbol{D}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{D}')\right) \tag{2.3.5}$$

$$\begin{aligned}\frac{\partial^2 \boldsymbol{Q}_2}{\partial \phi^2} =& -\frac{1}{2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial \boldsymbol{H}}{\partial \phi}\boldsymbol{H}^{-1}\frac{\partial \boldsymbol{H}}{\partial \phi}\right) - \frac{1}{2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial^2 \boldsymbol{H}}{\partial \phi^2}\right)\\ & - \frac{1}{2\sigma^2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial^2 \boldsymbol{H}}{\partial \phi^2}\boldsymbol{H}^{-1}(\boldsymbol{D}+\boldsymbol{D}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{D}')\right)\\ & + \frac{1}{\sigma^2}\mathbf{tr}\left(\boldsymbol{H}^{-1}\frac{\partial \boldsymbol{H}}{\partial \phi}\boldsymbol{H}^{-1}\frac{\partial \boldsymbol{H}}{\partial \phi}\boldsymbol{H}^{-1}(\boldsymbol{D}+\boldsymbol{D}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{D}')\right)\end{aligned} \tag{2.3.6}$$

here,

$$\left(\frac{\partial \boldsymbol{H}}{\partial \phi}\right)_{ij} = [1 - I_0(|s_i - s_j|)]\mathbf{exp}(-\phi|s_i - s_j|)(-|s_i - s_j|);$$

$$\left(\frac{\partial^2 \boldsymbol{H}}{\partial \phi^2}\right)_{ij} = [1 - I_0(|s_i - s_j|)]\mathbf{exp}(-2\phi|s_i - s_j|)(-|s_i - s_j|)^2,$$

$I_0(|s_i - s_j|) = 1$ if $|s_i - s_j| = 0,$ and 0 otherwise; $i, j = 1, 2, \cdots, n$.

2.4 Accelerated EM: The PX-EM Algorith

See Liu C, Rubin D B, Wu Y N(1998), we have

E-step:

$$Q_1 = -\frac{n}{2}\log \tau^2 - \frac{E_{\boldsymbol{w}|\boldsymbol{y},\boldsymbol{\theta}^{(t)}}(\boldsymbol{y} - \boldsymbol{X\beta} - \alpha\boldsymbol{w})'(\boldsymbol{y} - \boldsymbol{X\beta} - \alpha\boldsymbol{w})}{2\tau^2}$$

$$Q_2 = -\frac{n}{2}\log \sigma^2 - \frac{1}{2}\log|\boldsymbol{H}(\phi)| - \frac{\alpha^2 E[\boldsymbol{w'}\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}]}{2\sigma^2}.$$

M-step: Therefore, to maximize Q_1 and Q_2 and then obtain

$$\hat{\tau}^2 = \frac{(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) - 2\alpha(\boldsymbol{y} - \boldsymbol{X\beta})'E(\boldsymbol{w}) + \alpha^2 E(\boldsymbol{w'w})}{n} \tag{2.4.1}$$

$$\hat{\sigma}^2 = \frac{\alpha^2 E[\boldsymbol{w'}\boldsymbol{H}^{-1}(\phi)\boldsymbol{w}]}{n} \tag{2.4.2}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'}(\boldsymbol{y} - \alpha E(\boldsymbol{w})) \tag{2.4.3}$$

$$\hat{\alpha} = (\boldsymbol{y} - \boldsymbol{X\beta})'E(\boldsymbol{w})/E(\boldsymbol{w'w}) \tag{2.4.4}$$

3 Appendix A

3.1 VB

Because of

$$p\left(\tau^2|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta},\boldsymbol{w}\right)=IG\left(a_\tau+\frac{n}{2},b_\tau+\frac{(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})}{2}\right), \tag{3.1.1}$$

hence, only need to consider the expectation of this term, $(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})$, which is

$$\begin{aligned} E_{\beta,w}[(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta}-\boldsymbol{w})] =&(\boldsymbol{y}-\boldsymbol{\mu}_w)'(\boldsymbol{y}-\boldsymbol{\mu}_w)+\mathbf{tr}(\boldsymbol{\Sigma}_w)- \\ &2(\boldsymbol{y}-\boldsymbol{\mu}_w)'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{\mu}_w)+ \\ &E_\beta(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}) \end{aligned}$$

here,

$$\begin{aligned} E_\beta(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}) &= E_\beta[\mathbf{tr}(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta})] = E_\beta[\mathbf{tr}(\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}')] \\ &= \mathbf{tr}\{\boldsymbol{X}E_\beta[\boldsymbol{\beta}\boldsymbol{\beta}']\boldsymbol{X}'\} \\ &= \mathbf{tr}\{\boldsymbol{X}(\boldsymbol{V}_\beta+\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta')\boldsymbol{X}'\} \end{aligned}$$

Especially, $E_\beta(\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}) = pE(\frac{1}{\tau^2}) + (\boldsymbol{y}-\boldsymbol{\mu}_w)'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{y}-\boldsymbol{\mu}_w)$, as $\boldsymbol{\Sigma}_\beta=\mathbf{0}$.

4 Data fusion

Let $\{\phi_{ik}; i=1,2, k=1,2,\cdots,r.\}$ be pre-specified spatial basis functions, then we have

$$Z(s)=\begin{pmatrix} X(s) \\ Y(s) \end{pmatrix} = \begin{pmatrix} W(s) \\ W(s) \end{pmatrix} + \begin{pmatrix} \Sigma_{k=1}^r \phi_k(s)\eta_k \\ 0 \end{pmatrix} + \begin{pmatrix} \xi_1(s) \\ \xi_2(s) \end{pmatrix} \tag{4.0.2}$$