递推线性最小均方估计

基于矢量空间的递推线性最小均方估计算法推导

- 问题引入
- 基于矢量空间的推导
- 推导过程的几何演示
- 拓展：正交投影性质的补充
- 推导过程总结

1. 问题引入

考虑高斯白噪声中随机变量的估计问题，

$$z_i = A + w_i \qquad i = 0,1,\ldots,N-1$$

其中 w_i 是均值为零、方差为σ^2的高斯白声，$A \sim N(0,\sigma_A^2)$ ，且 w_i 与A统计独立。

根据N个观测得到的A的估计与均方误差为

$$\hat{A}[N-1] = \frac{\sigma_A^2}{\sigma_A^2 + \dfrac{\sigma^2}{N}}\bar{z} \qquad Mse(\hat{A}[N-1]) = \frac{\sigma_A^2\sigma^2}{N\sigma_A^2 + \sigma^2}$$

1. 问题引入

经过简单数学推导，可得

$$\hat{A}[N] = \hat{A}[N-1] + K[N](z_N - \hat{A}[N-1])$$

$$K[N] = \frac{Mse(\hat{A}[N-1])}{Mse(\hat{A}[N-1]) + \sigma^2}$$

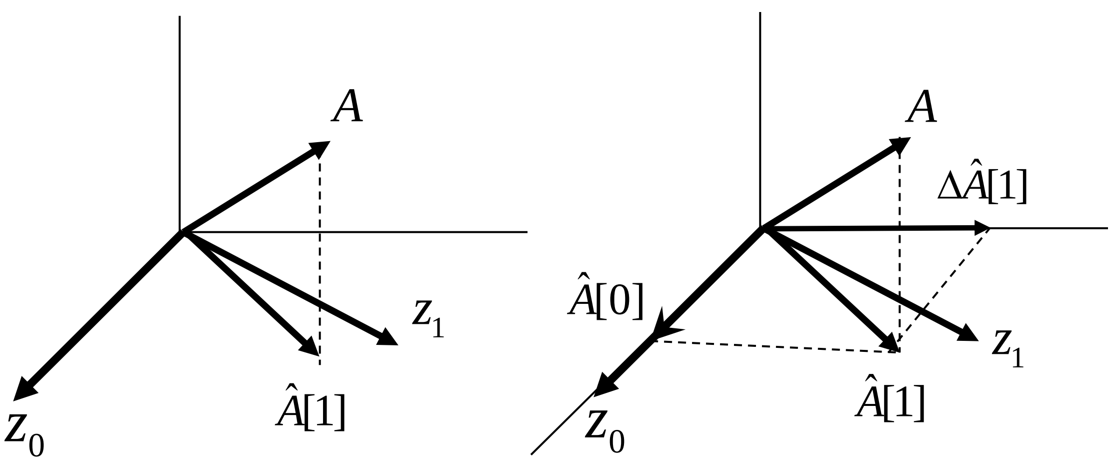
$$Mse(\hat{A}[N]) = (1 - K[N])Mse(\hat{A}[N-1])$$

教材上有详细的推导

2. 随机矢量空间的方法推导递推算法

假定 $\hat{A}[1]$ 是基于前两个观测数据 z_0, z_1 估计，

$\hat{A}[1]$ 可以看作为 $\hat{A}[0]$ 及一个与 $\hat{A}[0]$ 正交的矢量和。



$$\hat{A}[1] = \hat{A}[0] + \Delta\hat{A}[1]$$

$$\begin{aligned}
\hat{A}[0] &= \hat{E}(A \mid z_0) \\
&= \left(A, \frac{z_0}{\| z_0 \|}\right) \frac{z_0}{\| z_0 \|} \\
&= \frac{E(Az_0)}{E(z_0^2)} z_0 \\
&= \frac{E[A(A + w_0)]}{E(A^2) + E(w_0^2)} z_0 \\
&= \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2} z_0
\end{aligned}$$

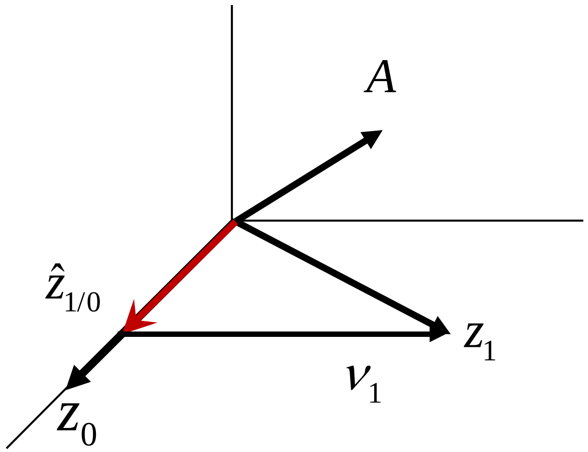
2. 随机矢量空间的方法推导递推算法

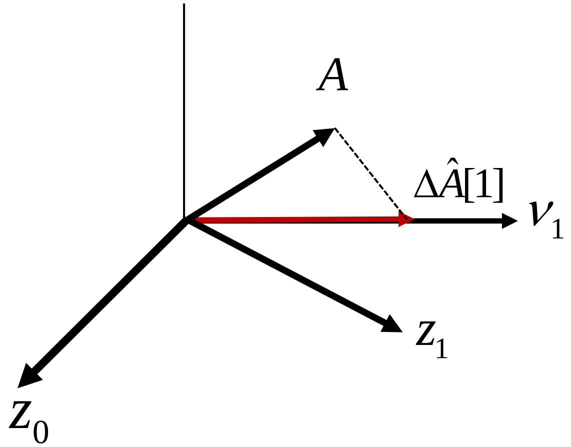
如何求得 $\Delta\hat{A}[1]$？

（1）根据 z_0 求 z_1 的估计 $\hat{z}_{1/0}$，得到的误差矢量 $v_1 = z_1 - \hat{z}_{1/0}$ 与 z_0 正交。

$$\hat{z}_{1/0} = \hat{E}(z_1 \mid z_0)$$

（2）将A投影到 $v_1 = z_1 - \hat{z}_{1/0}$ 上

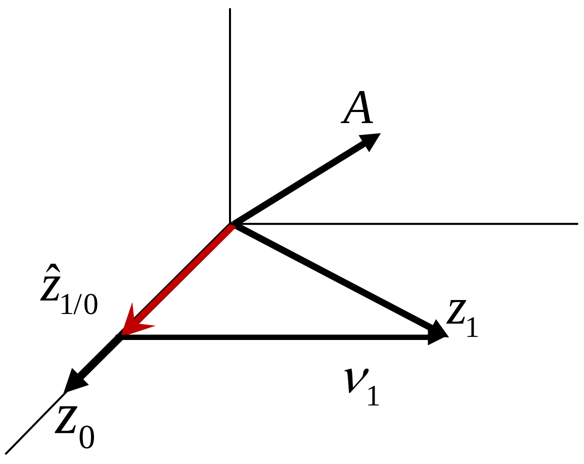




2. 随机矢量空间的方法推导递推算法

$$\hat{z}_{1/0} = \left(z_1, \frac{z_0}{\| z_0 \|} \right) \frac{z_0}{\| z_0 \|}$$

$$= \frac{(z_1, z_0)}{\| z_0 \|^2} z_0 = \frac{E(z_1 z_0)}{E(z_0^2)} z_0$$



$$\hat{z}_{1/0} = \frac{E[(A+w_1)(A+w_0)]}{E(A^2)+E(w_0^2)} z_0$$

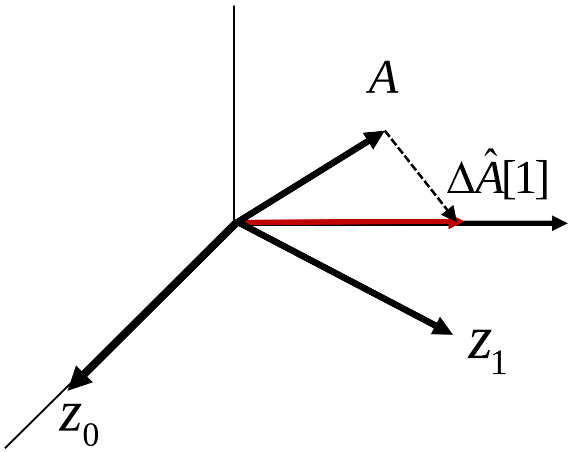
$$= \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2} z_0$$

$$v_1 = z_1 - \hat{z}_{1/0} = z_1 - \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2} z_0$$

2. 随机矢量空间的方法推导递推算法

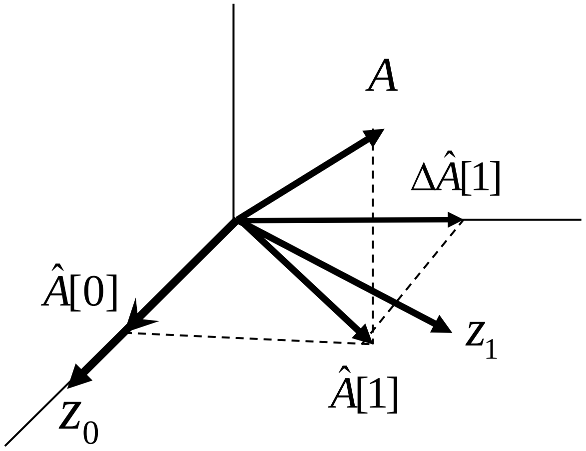
$$\Delta\hat{A}[1]=\left(A,\frac{v_1}{\|v_1\|}\right)\frac{v_1}{\|v_1\|}$$

$$=\frac{(A,v_1)}{\|v_1\|^2}v_1=\frac{E(Av_1)}{E(v_1^2)}v_1$$

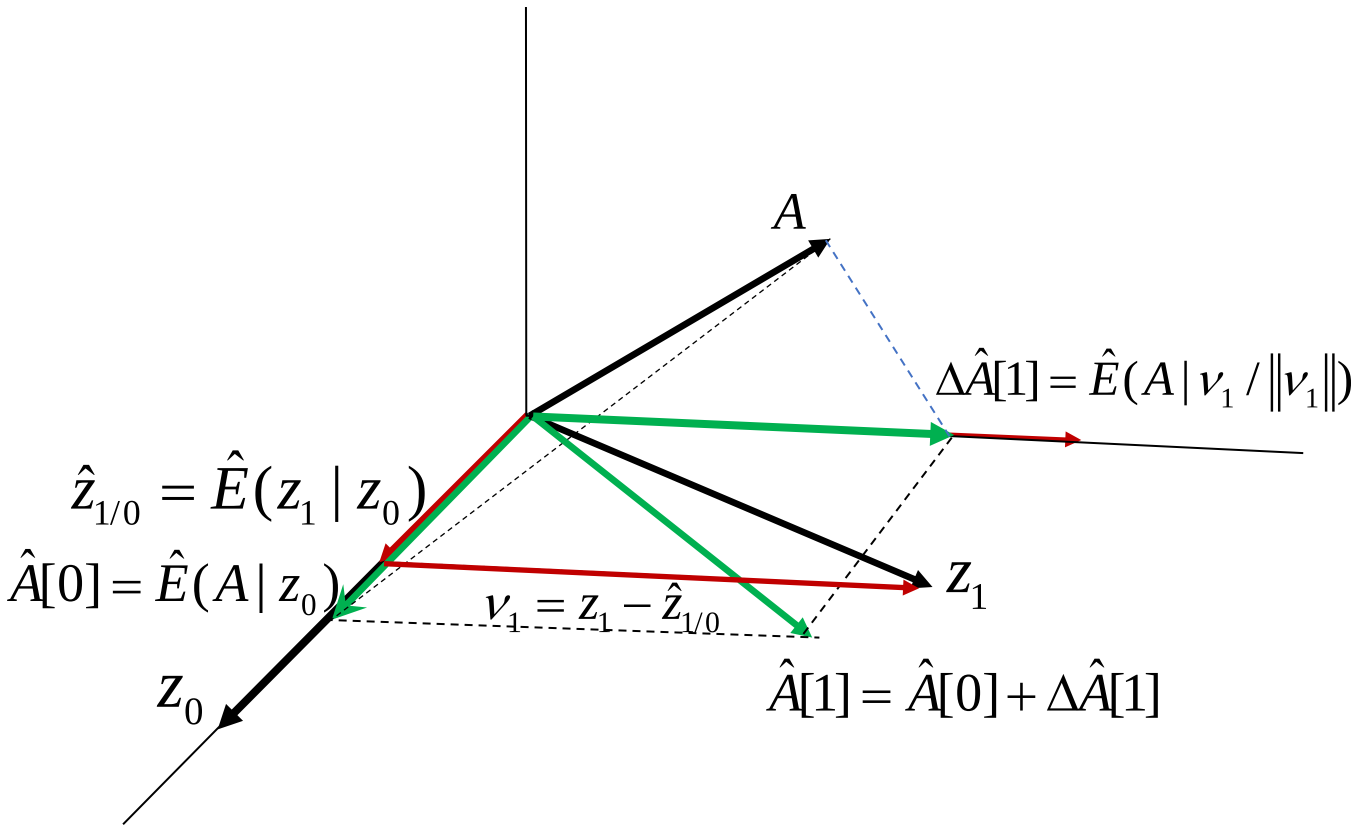


$$\hat{A}[1]=\hat{A}[0]+\frac{E(Av_1)}{E(v_1^2)}v_1=\hat{A}[0]+K[1](z_1-\hat{z}_{1/0})$$

$$K[1]=\frac{E(A\nu_1)}{E(\nu_1^2)}=\frac{E\left[A(z_1-\frac{\sigma_A^2}{\sigma_A^2+\sigma^2}z_0)\right]}{E\left[\left(z_1-\frac{\sigma_A^2}{\sigma_A^2+\sigma^2}z_0\right)^2\right]}=\frac{\sigma_A^2}{2\sigma_A^2+\sigma^2}$$



3.推导过程的几何解释

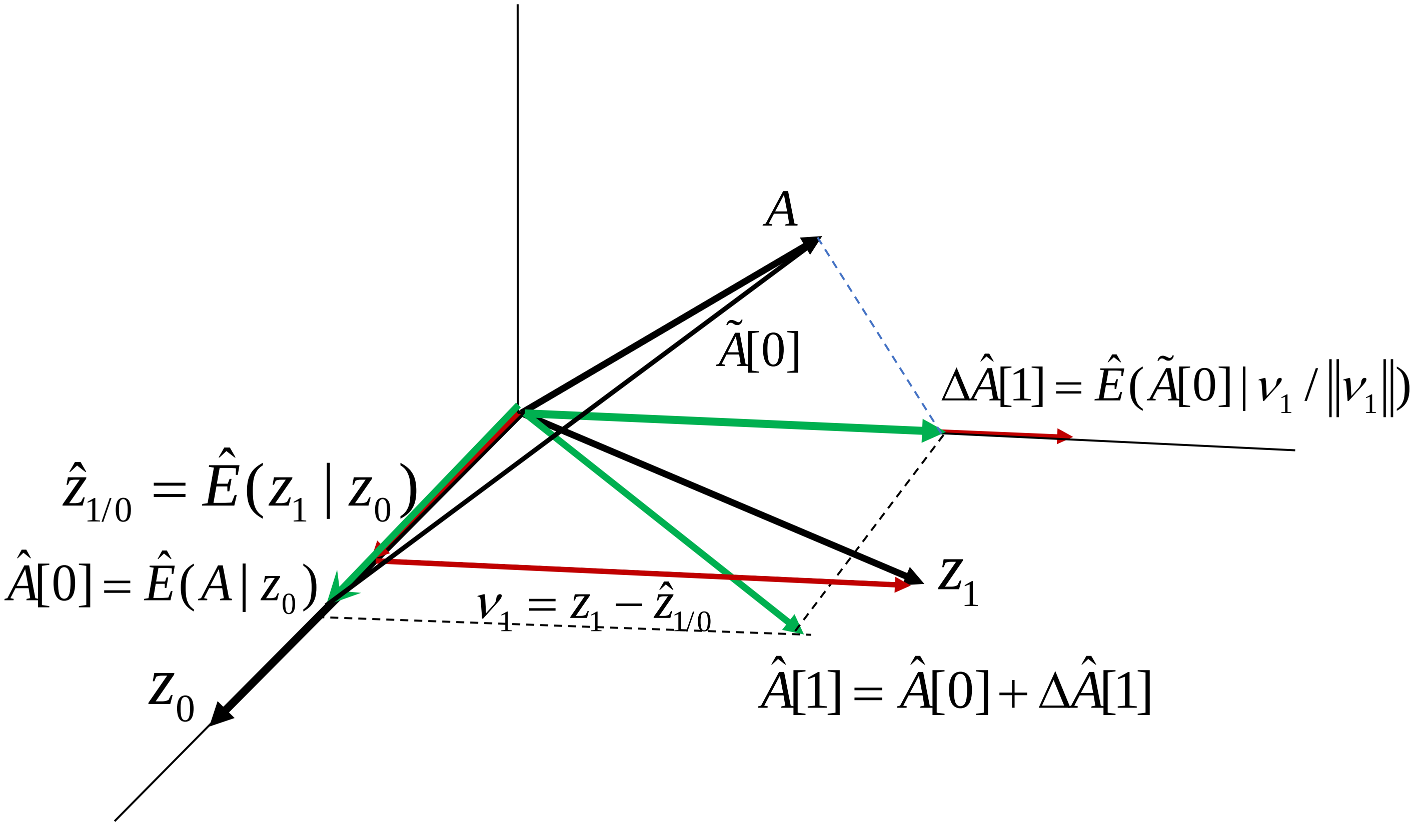


A
$\Delta\hat{A}[1]=\hat{E}(A\,|\,\nu_1/\|\nu_1\|)$
$\hat{z}_{1/0}=\hat{E}(z_1\,|\,z_0)$
$\hat{A}[0]=\hat{E}(A\,|\,z_0)$
z_1
$\nu_1=z_1-\hat{z}_{1/0}$
z_0
$\hat{A}[1]=\hat{A}[0]+\Delta\hat{A}[1]$

另一种形式

$$\Delta\hat{A}[1]=\hat{E}(A\,|\,v_1/\|v_1\|)=\hat{E}[(A-\hat{A}[0])\,|\,v_1/\|v_1\|)=\hat{E}(\tilde{A}[0]\,|\,v_1/\|v_1\|)$$

$$\hat{E}[\hat{A}[0]\,|\,v_1/\|v_1\|)=\frac{E(\hat{A}[0]v_1)v_1}{\|v_1\|^2}=0$$



4. 拓展：正交投影的性质（6）

（6）设 $\mathbf{z}^k = \begin{bmatrix} \mathbf{z}^{k-1} \\ \mathbf{z}_k \end{bmatrix} = \begin{bmatrix} \mathbf{z}_0 & \mathbf{z}_1 & \dots & \mathbf{z}_k \end{bmatrix}^T$

$$\hat{E}(\boldsymbol{\theta} \mid \mathbf{z}^k) = \hat{E}(\boldsymbol{\theta} \mid \mathbf{z}^{k-1}) + \hat{E}(\tilde{\boldsymbol{\theta}}[k-1] \mid \mathbf{v}[k])$$

$$= \hat{E}(\boldsymbol{\theta} \mid \mathbf{z}^{k-1}) + E(\tilde{\boldsymbol{\theta}}[k-1]\mathbf{v}^T[k])\{E(\mathbf{v}[k]\mathbf{v}^T[k])\}^{-1}\mathbf{v}[k]$$

其中 $\tilde{\boldsymbol{\theta}}[k-1] = \boldsymbol{\theta} - \hat{E}(\boldsymbol{\theta} \mid \mathbf{z}^{k-1})$

$$\mathbf{v}[k] = \mathbf{z}_k - \hat{E}(\mathbf{z}_k \mid \mathbf{z}^{k-1})$$

称为新息(Innovation)

5.推导过程总结

$$\hat{A}[1]=\hat{A}[0]+\Delta\hat{A}[1]$$

$$\hat{A}[0]=\frac{E(Az_0)}{E(z_0^2)}z_0=\frac{\sigma_A^2}{\sigma_A^2+\sigma^2}z_0$$

$$\hat{z}_{1/0}=\frac{E(z_1 z_0)}{E(z_0^2)}z_0=\frac{\sigma_A^2}{\sigma_A^2+\sigma^2}z_0$$

$$\nu_1=z_1-\hat{z}_{1/0}$$

$$\Delta\hat{A}[1]=\frac{E(A\nu_1)}{E(\nu_1^2)}\nu_1$$

或

$$\Delta\hat{A}[1]=\frac{E(\tilde{A}[0]\nu_1)}{E(\nu_1^2)}\nu_1$$

$$K[1]=\frac{E(A\nu_1)}{E(\nu_1^2)}=\frac{E(\tilde{A}[0]\nu_1)}{E(\nu_1^2)}=\frac{\sigma_A^2}{2\sigma_A^2+\sigma^2}$$

$$\hat{A}[1]=\hat{A}[0]+K[1](z_1-\hat{z}_{1/0})$$

按这个过程继续求 $\hat{A}[2]$

$$\hat{A}[2]=\hat{A}[1]+K[2](z_2-\hat{z}_{2/1})$$

5.推导过程总结

在这个过程中，我们获得一组相互正交的数据

$$z_0, z_1-\hat{z}_{1/0}, z_2-\hat{z}_{2/1}, \ldots.$$
$$\nu_0, \quad \nu_1, \quad \nu_2, \ldots.$$

根据LMMSE对于正交数据的相加性原理，

$$\hat{A}[N]=\sum_{i=0}^{N}\left(A,\frac{\nu_i}{\|\nu_i\|}\right)\frac{\nu_i}{\|\nu_i\|}=\sum_{i=0}^{N}\frac{(A,\nu_i)}{\|\nu_i\|^2}\nu_i=\sum_{i=0}^{N}\frac{E(A\nu_i)}{E(\nu_i^2)}\nu_i=\sum_{i=0}^{N}K[i]\nu_i$$

$$\hat{A}[N]=\sum_{i=0}^{N-1}K[i]\nu_i+K[N]\nu_N=\hat{A}[N-1]+K[N](z_N-\hat{z}_{N/N-1})$$

$$K[N]=\frac{E[A(z_N-\hat{z}_{N/N-1})]}{E[(z_N-\hat{z}_{N/N-1})^2]}=\frac{E[\tilde{A}[N-1](z_N-\hat{z}_{N/N-1})]}{E[(z_N-\hat{z}_{N/N-1})^2]}$$

5.推导过程总结

$$\hat{z}_{N/N-1} = \hat{E}((A+w_N)\mid z_0,\ldots.,z_{N-1}) = \hat{A}[N-1]+\hat{w}_{N/N-1} = \hat{A}[N-1]$$

$$z_N - \hat{z}_{N/N-1} = A + w_N - \hat{A}[N-1] = \tilde{A}[N-1] + w_N$$

$$E[(z_N - \hat{z}_{N/N-1})^2] = Mse(\hat{A}[N-1]) + \sigma^2$$

$$E[\tilde{A}[N-1](z_N - \hat{z}_{N/N-1})] = E\left\{\tilde{A}[N-1](\tilde{A}[N-1]+w_N)\right\}$$

$$\Downarrow$$

$$Mse(\hat{A}[N-1])$$

5.推导过程总结

$$\tilde{A}[N-1]+w_N$$

$$K[N]=\frac{E[\tilde{A}[N-1]\boxed{(z_N-\hat{z}_{N/N-1})}]}{E[(z_N-\hat{z}_{N/N-1})^2]}=\frac{Mse(\hat{A}[N-1])}{Mse(\hat{A}[N-1])+\sigma^2}$$

接下来讨论$Mse(\hat{A}[N])$的计算：

$$\begin{aligned}A-\hat{A}[N]&=A-\hat{A}[N-1]-K[N](z_N-\hat{z}_{N/N-1})\\&=A-\hat{A}(N-1)-K[N](\tilde{A}[N-1]+w_N)\\&=(1-K(N))\tilde{A}[N-1]-K[N]w_N\end{aligned}$$

即 $$\tilde{A}[N]=(1-K(N))\tilde{A}[N-1]-K[N]w_N$$

5.推导过程总结

$$\tilde{A}[N]=(1-K(N))\tilde{A}[N-1]-K[N]w_N$$

$$Mse(\hat{A}[N])=E\left\{\left[(1-K(N))\tilde{A}[N-1]-K[N]w_N\right]^2\right\}$$

$$=(1-K(N))^2Mse(\hat{A}[N-1])+K^2[N]\sigma^2$$

$$=(1-2K(N))Mse(\hat{A}[N-1])+K^2(N)Mse(\hat{A}[N-1])+K^2[N]\sigma^2$$

$$=(1-2K(N))Mse(\hat{A}[N-1])+K^2\underbrace{(N)\left\{Mse(\hat{A}[N-1])+\sigma^2\right\}}_{K[N]Mse(\hat{A}[N-1])}$$

$$Mse(\hat{A}[N])=(1-K[N])Mse(\hat{A}[N-1])$$

$$K[N]=\frac{Mse(\hat{A}[N-1])}{Mse(\hat{A}[N-1])+\sigma^2}$$

5.推导过程总结

估计器更新： $\hat{A}[N]=\hat{A}[N-1]+K[N](z_N-\hat{z}_{N/N-1})$

$$\hat{z}_{N/N-1}=\hat{A}[N-1]$$

增益计算： $K[N]=\dfrac{Mse(\hat{A}[N-1])}{Mse(\hat{A}[N-1])+\sigma^2}$

均方误差更新： $Mse(\hat{A}[N])=(1-K[N])Mse(\hat{A}[N-1])$

起始条件： $\hat{A}[0]=\dfrac{E(Az_0)}{E(z_0^2)}z_0=\dfrac{\sigma_A^2}{\sigma_A^2+\sigma^2}z_0$

$$Mse(\hat{A}[0])=\frac{\sigma_A^2\sigma^2}{\sigma_A^2+\sigma^2}$$