$\Rightarrow \left\{ R \in \mathcal{L}_2(Y,X) \Rightarrow \mathrm{Cov}(RZ) \in \mathcal{L}_1(X) \text{ for any } Y\text{-HSP } Z \right\}$

constraint for reconstruction

$$T^+ g = \sum_j \sigma_j^{-1} (g_1, g_2)_Y f_j \quad \text{with sing. system } \{\sigma_j, f_j, g_j\} \text{ for } T$$

General spectral reconstruction

$$Rg = \sum_j \frac{\lambda_j}{\sigma_j} (g_1, g_2) f_j \in \mathcal{L}_2, \text{ if } \sum_j \left(\lambda_j/\sigma_j\right)^2 < \infty$$

weights

$\Rightarrow \lambda_j$ has to decay much faster than σ_j

e.g. Tikhonov

$$Rg = \sum_j \frac{1}{\alpha + \sigma_j^2} (T^* g, f_j) f_j = \sum_j \frac{\sigma_j}{\alpha + \sigma_j^2} (g, g_j) f_j$$

$R \in \mathcal{L}_2$ if $\sum_j \left(\frac{\sigma_j}{\alpha + \sigma_j^2}\right)^2 < \infty$, which is satisfied if $T \in \mathcal{L}_2(X, Y)$

5§1.3 The minimax linear estimator

$\rightarrow$ Setting from above

we could also write $\mathcal{L}_2(Y, \ldots$

<u>Def 5§1.3.B</u> Risk of estimator $R \in \mathcal{L}(Y, X)$ under a priori knowledge $f^+ \in K$, $K \subset X$, and noise model Z

$$\mathrm{risk}_R(Z, K) := \sup_{f \in K} \| R(Tf + Z) - f \|^2$$

Implicit assumption: $\mathrm{Cov}(RZ) \in \mathcal{L}_1(X)$

$\rightsquigarrow \| R(Tf + Z) - f \|^2 = \mathbb{E}(\| R(Tf + Z) - f \|_X^2)$

Bias variance decomposition

$$\mathbb{E}\left(\| R(Tf + Z) - f \|_X^2\right) = \| (RT - \mathrm{Id}) f \|_X^2 + \mathbb{E}\left(\| RZ \|_X^2\right) \qquad (1.5.2.F)$$

<u>Def 5§1.3.C:</u> minimax linear risk

$$\mathrm{mlr}(Z, K) = \inf_{\substack{R \in \mathcal{L}(Y,X) \\ \mathrm{cov}(RZ) \in \mathcal{L}_1}} \mathrm{risk}_R(Z, K)$$

Special: (i) $Z = \varepsilon W$ $\quad \mathbb{E}(W) = 0 \quad \mathrm{Cov}(W) = \mathrm{Id}$

stochastic noise level

(ii) $K = K_{\nu, \rho} := \left\{ f \in X : \sum \sigma_j^{-2\nu} (f, f_j)^2 \le \rho^2 \right\}$

$$R_{\{\lambda\}} g = \sum_{j=1}^{\infty} \frac{\lambda_j}{\sigma_j} (g, g_j) f_j \in \mathcal{L}_2(Y, X) \qquad (5.1.3.D)$$

$$\Rightarrow \text{risk}_{R\{\lambda\}}(\varepsilon W, K_{\nu,\rho}) \overset{\sup_{f\in K_{\nu,\rho}}}{=} \|(RT - Id) f\|_X^2 + \varepsilon^2 \|RW\|^2 =$$

$$\left[(RT - Id) f = \sum_j (\lambda_j - 1)(f, f_j) f_j \right]$$

$$\overset{\sup_{f\in K_{\nu,\rho}}}{=} \sum_j (\lambda_j - 1)^2 (f, f_j)^2 + \varepsilon^2 \underbrace{tr(RR^*)}_{= \sum_j (\lambda_j/\sigma_j)^2}$$

$$= \sup_{x_j: \sum x_j^2 \sigma_j^{-2\nu} \le \rho^2} \sum_j (1 - \lambda_j)^2 x_j^2 + \varepsilon^2 (\lambda_j / \sigma_j)^2$$

$\xi_j := \frac{x_j}{\rho} \sigma_j^{-\nu}$

$$= \sup_{\sum_j \xi_j^2 \le 1} \sum_j \left((1 - \lambda_j)^2 \sigma_j^{2\nu} \rho^2 \xi_j^2 + \varepsilon^2 (\lambda_j / \sigma_j)^2 \right)$$

$$\text{risk}_{R\{\lambda\}}(\varepsilon W, K_{\nu,\rho}) = \rho^2 \max_j \left\{ (1 - \lambda_j)^2 \sigma_j^{2\nu} \right\} + \varepsilon^2 \sum_{j=1}^{\infty} (\lambda_j / \sigma_j)^2 \qquad (5.1.3.F)$$

<u>Lemma 5.1.3 E</u>: If $Z = \varepsilon W$, then every minimax linear estimator (i.e., one that realizes $m\,e_r$) has the form (5.1.3.D)

<u>Proof</u>: $S \in \mathcal{L}_2(Y, X)$, Idea: Build $R_{\{\lambda\}}$, which is at least as good as S

Ingenious choice: $\lambda_j := \sigma_j (S g_j, f_j)_X$

Observation: let k be the index that yields the max in (5.1.3 F) for the given sequence $\{\lambda_j\}$

$f^* = \rho \sigma_k^{\nu} f_k$ is the $f \in K_{\nu,\rho}$ that realizes the sup in $\text{risk}_{R\{\lambda\}}(\varepsilon W, K_{\nu,\rho})$

$$\text{risk}_{R\{\lambda\}}(\varepsilon W, K_{\nu,\rho}) = \text{risk}_{R\{\lambda\}}(\varepsilon W, \{f^*\}) =$$

$$= \underbrace{\|(R_{\{\lambda\}} T - Id) f^*\|_X^2}_{= (\lambda_k - 1)^2 \rho^2 \sigma_k^{2\nu}} + \underbrace{\mathbb{E}(\|R_{\{\lambda\}} Z\|^2)}_{= \varepsilon^2 tr(R^*_{\{\lambda\}} R_{\{\lambda\}})} \qquad (*)$$

$$\Gamma(ST - Id)f^* \underset{Tf_k = \sigma_k g_k}{=} \sigma_k^{\nu+1} \varrho S g_k - f^*$$

$$\| (ST - Id) f^* \|_X^2 = \sum_{j \neq k} \sigma_k^{2\nu+2} \varrho^2 (S g_k, f_j)^2 + \sigma_k^{\nu+1} \varrho (S g_k, f_k)_X - \varrho \sigma_k^{\nu})^2$$

$$\geq \underline{\varrho^2 \sigma_k^{2\nu} (\lambda_k - 1)^2}$$

$$(*) \leq \| (ST - Id) f^* \|_X^2 + \varepsilon^2 \underbrace{\| R_{\{\lambda\}} \|_{\mathcal{J}_2}^2}$$

$$= \sum_j \| R_{\{\lambda\}} g_j \|_X^2 =$$

$$= \sum_j (\lambda_j / \sigma_j)^2 = \sum_j (S g_j, f_j)_X^2 = \| S \|_{\mathcal{L}_2}^2$$

$$= tr(Cov(SW))$$

$$\Rightarrow risk_{R\{\lambda\}} (\varepsilon W, K_{\nu,\varrho}) \leq \| (ST - Id) f^* \|_X^2 + \varepsilon^2 tr(Cov(SW))$$

$$= E(\| S(Tf^* + Z) - f^* \|_X^2)$$

$$\leq risk_S (\varepsilon W, K_{\nu,\varrho})$$

$\square$

4.12.2013

Recall: $risk_{R\{\lambda\}} (\varepsilon W, K_{\nu,\varrho}) = \underbrace{\varrho^2 \sup_j (1 - \lambda_j)^2 \sigma_j^{2\nu}}_{\sup_{f \in K_{\nu,\varrho}} \| (R_{\{\lambda\}} T - Id) f \|_X^2} + \underbrace{\varepsilon^2 \sum_j (\lambda_j / \sigma_j)^2}_{tr(Cov(R_{\{\lambda\}} Z))}$

Goal: Find λ_j s.t. risk becomes minimal
(For-the-sake-of-simplicity assumption: T injective)

(i) for any minimizing sequence: $0 \leq \lambda_j \leq 1$
$\Rightarrow \sup_j (1 - \lambda_j)^2 \sigma_j^{2\nu} \leq \sigma_1^{2\nu}$

(ii) $S_k := \{ (\lambda_j)_{j \in \mathbb{N}} : 0 \leq \lambda_j \leq 1, \sup_j |1 - \lambda_j| \sigma_j^{\nu} = k \}$, $0 \leq k \leq \sigma_1^{\nu}$

Best sequence in S_k: $\lambda_j^k := \max\{1 - k\sigma_j^{-\nu}, 0\}$

$[(1 - \lambda_j)\sigma_j^{\nu} = k \rightarrow \lambda_j = 1 - k\sigma_j^{\nu}$ (must not be < 0 !)

$risk_{R\{\lambda^k\}} (\varepsilon W, K_{\nu,\varrho}) = \varrho^2 k^2 + \varepsilon^2 \sum_{j=1}^{J(k)} \left(\frac{\lambda_j^k}{\sigma_j} \right)^2$ (note only finitely many λ_j^k are non-zero) (*)