

Dieses Kapitel ist noch seeeeehr work in progress also ich hab erstmal grob Gedanken aufgeschrieben.

0.1 paper [1]

Beim Messen unserer Daten entstehen immer Ungenauigkeiten.

0.2 Idee

Wir betrachten die gemessene Datenmatrix

$$X = X^{\text{true}} + \gamma X^{\text{noise}} \quad (0.1)$$

mit $X \in \mathbb{R}^{n \times m}$, X^{true} als *echte* Datenmatrix und *approximately of low rank* und X^{noise} als Messfehlermatrix welche Unabhängig und identisch verteilte Zufallsvariablen enthält *vielleicht noch definieren*.

Hier benutzen wir die *truncated* Singulärwertzerlegung also

$$X = \sum_{i=1}^n \sigma_i u_i v_i' . \quad (0.2)$$

Schätzen kann man die *truncation* jetzt mit

$$\hat{X}_r = \sum_{i=1}^r \sigma_i u_i v_i' . \quad (0.3)$$

mit $r = \text{rang}(X^{\text{true}})$ und den Singulärwerten wieder nach Größe absteigend sortiert.

Da aber der Rang von X^{true} nicht unbedingt bekannt ist kann man versuchen, anhand der Größe, der Singulärwerte zu Kürzen. Dafür bauen wir mithilfe einer Indikatorfunktion $\eta_H(\sigma_i; \tau) = \sigma_i \chi_{\{\sigma_i \geq \tau\}}$ die gekürzte Matrix

$$\hat{X}_\tau = \sum_{i=1}^r \eta_H(\sigma_i; \tau) u_i v_i' . \quad (0.4)$$

Jetzt gilt es aber herauszufinden wie wir dieses τ wählen damit wir immernoch sinnvolle Ergebnisse erhalten, aber gleichzeitig zu viel rechenaufwand spaaren wie möglich.

0.3 Setting

Definition 0.1 (denoiser). Ein *Denoiser* reduziert eine Matrix auf nur noch *relevante* Ränge mit der Form

$$\hat{X}: \sum_{i=1}^n \sigma_i u_i v_i' \rightarrow \sum_{i=1}^n \eta(\sigma_i; \lambda) u_i v_i' \quad (0.5)$$

Definition 0.2 (Mean Square Error (MSE)). Mit dem MSE wird der Fehler der beim *denoising* Prozess entsteht, durch

$$\left\| \hat{X}(X) - X^{\text{true}} \right\|_F^2 = \sum_{i,j} \left(\hat{X}(X)_{i,j} - X_{i,j}^{\text{true}} \right)^2 \quad (0.6)$$

quantifiziert.

Wir betrachten eine Folge an denoising Problemen, die immer größer werden.

$$X_k = X_k^{\text{true}} + \gamma X_k^{\text{noise}} \quad (0.7)$$

mit $X_k^{\text{true}}, X_k^{\text{noise}} \in M_{n_k, m}$.

Jedes X_k^{noise} erfüllt die Eigenschaften von X^{noise} . Gegeben sei ein fixer Vektor $\mathbf{x} \in \mathbb{R}^r$, mit $r > 0$, den Einträgen $\mathbf{x} = (x_1, \dots, x_r)$ und $x_i \geq x_{i+1} > 0$. Für alle k gilt

$$X_k^{\text{true}} = U_k \text{diag}(x_1, \dots, x_r, 0, \dots, 0) V_k' . \quad (0.8)$$

Definition 0.3 (Asymptotic Mean Square Error (AMSE)). lalala