Subspace Analysis:

Note: Sensitivity analysis isolates subsets of influential parameters but ineffective for subspaces that are not aligned with coordinate axes.

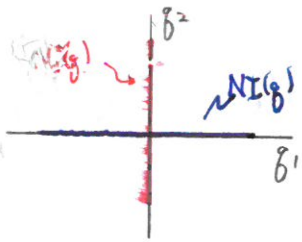
Linearly Parameterized Problems:

$$y = Aq \quad , y \in \mathbb{R}^n \;, q \in \mathbb{R}^p \;, \; A \text{ is } n \times p$$

e.g.) $y_i = q_2 x_i \quad , i = 1, 2, 3$

$q = [q_1, q_2]$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & x_1 \\ 0 & x_2 \\ 0 & x_3 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$



$$NI(q) = \mathcal{N}(A) = c\begin{bmatrix} 1 \\ 0 \end{bmatrix} \;, \; c \in \mathbb{R}$$ Null space of A

$; \; \mathcal{N}(A) = \{ q \in \mathbb{R}^p \mid Aq = 0 \}$

$$I(q) = \mathcal{R}(A^T) = c\begin{bmatrix} 0 \\ 1 \end{bmatrix} \;, \; c \in \mathbb{R}$$ Range

$\mathcal{R}(A^T) = \{ b \in \mathbb{R}^n \mid b = A^T z \text{ for some } z \in \mathbb{R}^n \}$

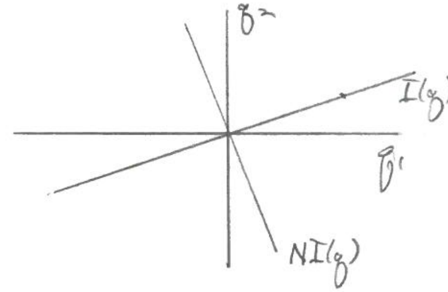
Note: $\mathcal{N}(\underbrace{A^T A}_{\text{Fisher information matrix}}) = \mathcal{N}(A) \;, \; \mathcal{R}(A^T A) = \mathcal{R}(A^T)$

Reference: Ilse C. F. Ipsen, "Numerical Matrix Analysis," SIAM 2009

e.g.) $y = [2 \;\; 1]\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$

$$NI(q) = \mathcal{N}(A) = c\begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix} \;, \; c \in \mathbb{R}$$ Not aligned with axes

$$I(q) = c\begin{bmatrix} 2 \\ 1 \end{bmatrix} \;, \; c \in \mathbb{R}$$



Deterministic Algorithms:

SVD: $$\underset{n \times p}{A} = \underset{n \times n}{U} \, \Sigma \, \underset{p \times p}{V^T}$$

U & V are orthogonal

. Left and right singular vectors

$$\underset{n \times p}{\Sigma} = [S \;,\; 0]$$

$$S = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix} \;, \; \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq \varepsilon$$

S_r

$U = [U_r \quad U_{n-r}]$

$V_r \in \mathbb{R}^{p \times r}, \; V_{p-r} \in \mathbb{R}^{p \times (p-r)}$

$U_r \in \mathbb{R}^{n \times r}$

$$V = [\underbrace{V_r}_{\text{Basis for } A^T} \quad \underbrace{V_{p-r}}_{\text{Basis for } \mathcal{N}(A)}]$$

Note: $A = \underset{n\times r}{U_r}\ \underset{r\times r}{S_r}\ \underset{r\times p}{V_r^T}$

$A^T \in \mathbb{R}^{p\times n}$, $p \geq n$

QR Algorithms: Focus on A^T – "Tall and skinny" and want $\mathcal{R}(A^T)$

$$\begin{bmatrix} A^T \end{bmatrix}_{p\times n} = \begin{bmatrix} Q \end{bmatrix}_{p\times p} \begin{bmatrix} R \\ \end{bmatrix}_{p\times n}$$

Upper Triangular

Rank-Revealing Algorithms:

$r\times r$

$$A^T P = QR = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

- Ensures that first r vectors of Q are basis for A^T.

Randomized Algorithms: [Halko, Martinsson, Tropp – SIAM Review 2011]

Goal: Find low-dimensional subspace that approximates action of A by sampling $y = Aq$. Find Q whose r-dimensional range approximates $\mathcal{R}(A)$ in sense

$$\|A - QQ^T A\| \leq \varepsilon$$

Algorithm: i) Choose ℓ random inputs q^i and compute $y^i = Aq^i$. Form

$$Y = \begin{bmatrix} y^1, \cdots, y^\ell \end{bmatrix}_{n\times \ell}$$

ii) $Y = QR$

Example: Take

$$y_i = \sum_{k=1}^{p} q_k \sin(2\pi k t_i)$$

evaluated at $t_i = (i-1)\Delta t$, $\Delta t = \frac{1}{n-1}$, $n = 1, \cdots, n$ on $[0,1]$.

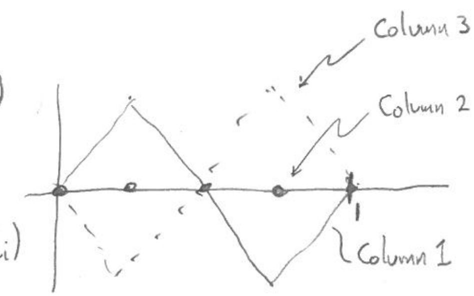
$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sin(2\pi t_1) & \sin(4\pi t_1) & \cdots & \sin(2\pi p t_1) \\ \vdots & & & \\ \sin(2\pi t_n) & & \cdots & \sin(2\pi p t_n) \end{bmatrix} \begin{bmatrix} q^1 \\ \vdots \\ q^p \end{bmatrix}$$

Case i: $p = n = 5$

Col 1: $\sin(2\pi t_1) \cdots \sin(2\pi t_n)$

Col 2: $\sin(4\pi t_1) \cdots \sin(4\pi t_n)$

Col 3: $\sin(2\pi \cdot 3 t_i) = -\sin(2\pi t_i)$

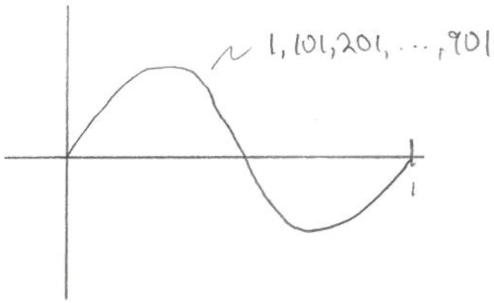


Basis for $I(q) = \mathcal{R}(A^T)$:

$$V_r(:,1) = Q(:,1) = \begin{bmatrix} -0.5774 \\ 0 \\ 0.5774 \\ 0 \\ -0.5774 \end{bmatrix}$$

Case ii: $n = 101$, $p = 1000$ Code: Example6-11.m

$r = 49$ - Depends on tolerance



Case iii: $n = 100$

$p = 10^6$

$r = 49$

Nonlinearly Parameterized Problems: $y = f(q)$

Example: See PowerPoint notes

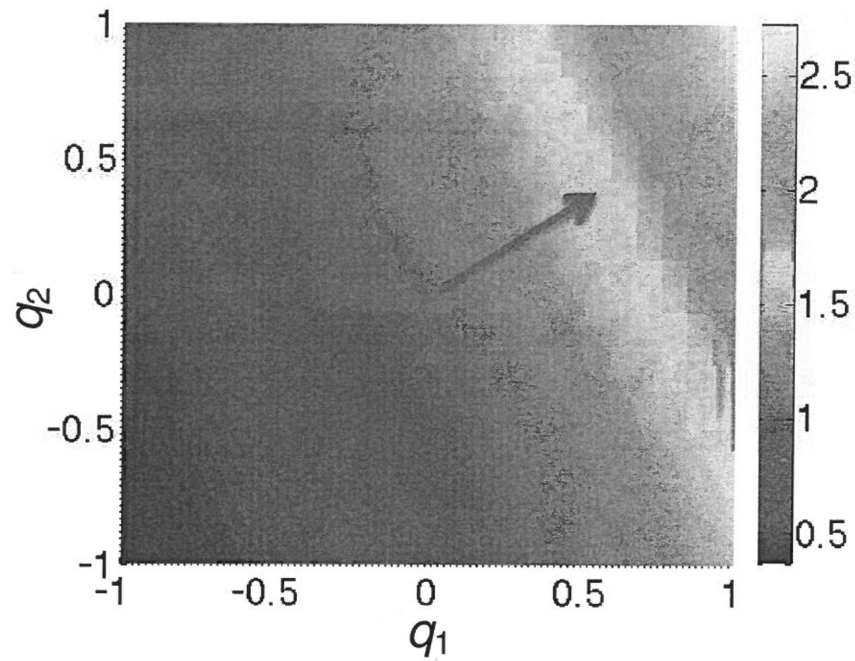
Active Subspaces

Note:

- Functions may vary significantly in only a few directions
- "Active" directions may be linear combination of inputs

Example: $y = \exp(0.7q_1 + 0.3q_2)$

- Varies most in [0.7, 0.3] direction
- No variation in orthogonal direction



A Bit of History:

- Often attributed to Russi (2010).
- Concept same as *identifiable subspaces* from systems and control; e.g., Reid (1977).
- For linearly parameterized problems, active subspace given by SVD or QR; Beltrami (1873), Jordan (1874), Sylvester (1889), Schmidt (1907), Weyl (1912). See 1993 *SIAM Review* paper by Stewart.

Gradient-Based Active Subspace Construction

Active Subspace: Consider

$$f = f(q)\ , q \in \mathcal{Q} \subseteq \mathbb{R}^p$$

- E.g., see [Constantine, SIAM, 2015; Stoyanov & Webster, *IJUQ*, 2015]

and

$$\nabla_q f(q) = \left[\frac{\partial f}{\partial q_1}, \cdots, \frac{\partial f}{\partial q_p}\right]^T$$

Construct outer product

$\rho(q)$: Distribution of input parameters q

$$C = \int (\nabla_q f)(\nabla_q f)^T \rho dq$$

Question: How sensitive are results to distribution, which is typically not known?

Partition eigenvalues: $C = W \Lambda W^T$

$$\Lambda = \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix}, \ W = [W_1 \quad W_2]$$

Rotated Coordinates:

$$y = W_1^T q \in \mathbb{R}^n \quad \text{and} \quad z = W_2^T q \in \mathbb{R}^{p-n}$$

Active Variables

Active Subspace: Range of eigenvectors in W_1

Gradient-Based Active Subspace Construction

Active Subspace: Construction based on random sampling

1. Draw M independent samples $\{q^j\}$ from ρ
2. Evaluate $\nabla_q f_j = \nabla_q f(q^j)$
3. Approximate outer product

$$C \approx \widetilde{C} = \frac{1}{M}\sum_{j=1}^{M} (\nabla_q f_j)(\nabla_q f_j)^T$$

   Note: $\widetilde{C} = GG^T$ where $G = \frac{1}{\sqrt{M}}[\nabla_q f_1, \ldots, \nabla_q f_M]$

4. Take SVD of $G = W\sqrt{\Lambda}V^T$
   - Active subspace of dimension n is first n columns of W

One Goal: Develop efficient algorithm for codes that do not have adjoint capabilities

Note: Finite-difference approximations tempting but not effective for high-D

Strategy: Algorithm based on initialized adaptive Morris indices

Example: Consider

$$y = e^{c_1 q_1 + c_2 q_2} = f(q)$$

so

$$\nabla_q f(q) = \begin{bmatrix} c_1 e^{c_1 q_1 + c_2 q_2} \\ c_2 e^{c_1 q_1 + c_2 q_2} \end{bmatrix} = \begin{bmatrix} c_1 f(q) \\ c_2 f(q) \end{bmatrix}$$

For $Q_1, Q_2 \sim \mathcal{U}(0,1)$, we have

$$C = \int_0^1 \int_0^1 (\nabla_q f)(\nabla_q f)^T dq_1 dq_2$$

$$= \int_0^1 \int_0^1 \begin{bmatrix} c_1^2 f^2(q) & c_1 c_2 f^2(q) \\ c_1 c_2 f^2(q) & c_2^2 f^2(q) \end{bmatrix} dq$$

$$= \begin{bmatrix} c_1^2 & c_1 c_2 \\ c_1 c_2 & c_2^2 \end{bmatrix} \cdot \frac{1}{4 c_1 c_2} (e^{2c_1} - 1)(e^{2c_2} - 1)$$

$$= \begin{bmatrix} \frac{c_1}{4c_2} & \frac{1}{4} \\ \frac{1}{4} & \frac{c_2}{4c_1} \end{bmatrix} (e^{2c_1} - 1)(e^{2c_2} - 1)$$

$\boxed{\text{e.g.}}$ $c_1 = 0.7$

$c_2 = 0.3$

Analytic C:

$$C = \begin{bmatrix} 1.4652 & 0.6279 \\ 0.6279 & 0.2691 \end{bmatrix}$$

Monte Carlo Approximation:

$$C \approx \frac{1}{M} \sum_{j=1}^{M} (\nabla_q f(q^j))(\nabla_q f(q^j))^T$$

$M = 10^4$	$M = 10^5$	$M = 10^6$
$\begin{bmatrix} 1.4532 & 0.6228 \\ 0.6228 & 0.2669 \end{bmatrix}$	$\begin{bmatrix} 1.4650 & 0.6278 \\ 0.6278 & 0.2691 \end{bmatrix}$	$\begin{bmatrix} 1.4654 & 0.6280 \\ 0.6280 & 0.2692 \end{bmatrix}$

Results:

```
>> [V1,D] = eig(C);
```

$$D = \begin{bmatrix} 1.7343 & 0 \\ 0 & 0 \end{bmatrix}$$ so 1-D subspace

$$V1 = \begin{bmatrix} 0.91 & .39 \\ 0.39 & .91 \end{bmatrix}$$

Basis for subspace

```
>> [U,S,V] = svd(G);
```

$$Sv = \begin{bmatrix} 1.3130 & 0 \\ 0 & 0 \end{bmatrix}$$ $1.3130^2 = 1.72$

$$U = \begin{bmatrix} .91 & .39 \\ .39 & .91 \end{bmatrix}$$