Math Review

M. Sami Fadali
Professor of Electrical Engineering
Outline

- Matrix identities.
- Quadratic forms.
- Vector Calculus.
- Pseudoinverse.
- Trace of a matrix and its properties.
- Partitioned matrices.
Important Matrix Identities

- Inverse
  \[(A_1 A_2 \ldots A_n)^{-1} = A_n^{-1} A_{n-1}^{-1} \ldots A_1^{-1}\]

- Transpose
  \[(A_1 A_2 \ldots A_n)^T = A_n^T A_{n-1}^T \ldots A_1^T\]

- Inverse Transpose
  \[(A^{-1})^T = (A^T)^{-1}\]
Partitioned Matrix

\[ A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \]

\[ \det[A] = \begin{cases} 
\det(A_1) \det(A_4 - A_3 A_1^{-1} A_2), & A_1^{-1} \text{ exists} \\
\det(A_4) \det(A_1 - A_2 A_4^{-1} A_3), & A_4^{-1} \text{ exists} 
\end{cases} \]

\[ A^{-1} = \begin{bmatrix} 
(A_1 - A_2 A_4^{-1} A_3)^{-1} & -A_1^{-1} A_2 (A_4 - A_3 A_1^{-1} A_2) \\
-A_4^{-1} A_3 (A_1 - A_2 A_4^{-1} A_3)^{-1} & (A_4 - A_3 A_1^{-1} A_2) 
\end{bmatrix} \]
Quadratic Forms

\[ x^T Q x = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j \]

\[ = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{q_{ij} + q_{ji}}{2} x_i x_j + \sum_{i=1}^{n} q_{ii} x_i^2 \]

\[ = x^T \left( \frac{Q + Q^T}{2} \right) x \]

○ Assume symmetric matrix w.l.o.g.
Definite Matrices

- Positive Definite $x^T Q x > 0, \forall x \neq 0$
- Positive Semidefinite $x^T Q x \geq 0, \forall x \neq 0$
- Negative Definite $x^T Q x < 0, \forall x \neq 0$
- Negative Semidefinite $x^T Q x \leq 0, \forall x \neq 0$
Eigenvalues of Definite Matrices

- Positive Definite
  \[ \lambda_i > 0, \; i = 1, \ldots, n \]
- Positive Semidefinite
  \[ \lambda_i \geq 0, \; i = 1, \ldots, n \]
- Negative Definite
  \[ \lambda_i < 0, \; i = 1, \ldots, n \]
- Negative Semidefinite
  \[ \lambda_i \leq 0, \; i = 1, \ldots, n \]
Gradient Vector

\[ y = a^T x = x^T a = \sum_{i=1}^{n} a_i x_i \]

\[
\begin{bmatrix}
\frac{\partial y}{\partial x}
\end{bmatrix}^T
= \begin{bmatrix}
\frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \cdots & \frac{\partial y}{\partial x_n}
\end{bmatrix}
= \begin{bmatrix}
a_1 & a_2 & \cdots & a_n
\end{bmatrix}
= a^T
\]
Jacobian Matrix

\[
y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = Qx = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} x = \begin{bmatrix} q_1^T x \\ q_2^T x \\ \vdots \\ q_n^T x \end{bmatrix}
\]

\[
\frac{\partial y}{\partial x} = \begin{bmatrix} (\partial y_1/\partial x)^T \\ (\partial y_2/\partial x)^T \\ \vdots \\ (\partial y_n/\partial x)^T \end{bmatrix} = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} = Q
\]
Gradient of a Quadratic Form

- Recall
  \[ y = a^T x = x^T a, \quad \frac{\partial y}{\partial x} = a \]

- Apply to the quadratic form
  \[ y = x^T Q x = (Q^T x)^T x \]
  \[ \frac{\partial y}{\partial x} = (Q + Q^T)x \]
  \[ = 2Qx \text{ for } Q \text{ symmetric} \]
Minimization/Maximization

\[ J(x) = J(x_0) + \left( \frac{\partial J(x)}{\partial x} \right)^T \Delta x + \frac{1}{2!} \Delta x^T \left( \frac{\partial^2 J(x)}{\partial x^2} \right) \Delta x + O(\|\Delta x\|^3) \]

- Second-order approximation.
- **Necessary**: Zero gradient.
- **Sufficient for min**: Hessian positive definite.
- **Sufficient for max**: Hessian negative definite.
Singular Value Decomposition

- Decomposition for any matrix of rank \( r \).

\[
H = \underbrace{U}_{N \times n} \underbrace{\Sigma}_{N \times N} \underbrace{V^*}_{N \times n} n \times n
\]

\[
U^{-1} = U^* \quad V^{-1} = V^*
\]

\[
\Sigma = \begin{bmatrix}
\Sigma_r & 0_{r \times (n-r)} \\
0_{(N-r) \times r} & 0_{(N-r) \times (n-r)}
\end{bmatrix}
\]

\[
\Sigma_r = diag\{\sigma_1, \sigma_2, \ldots, \sigma_r\}
\]

\[
\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r
\]

- \( \sigma_i \) = singular values of the matrix \( H \).
Special Cases (full rank)

\[ H = U \Sigma V^* \]

\[ \begin{align*}
    r = N < n: & \quad \Sigma = [\Sigma_N | 0_{N \times (n-N)}] \\
    r = n < N: & \quad \Sigma = \begin{bmatrix} \Sigma_n \\ 0_{(N-n) \times n} \end{bmatrix} \\
    r = N = n: & \quad \Sigma = \Sigma_n
\end{align*} \]
Pseudoinverse

\[
H_{N \times n} = U_{N \times N} \Sigma_{N \times n} V^*_{n \times n}, \quad H^\# = V_{n \times N} \Sigma^\#_{n \times n} U^*_{N \times N}
\]

\[
\Sigma^\# = \begin{bmatrix}
\Sigma_r^{-1} & 0_{r \times (N-r)} \\
0_{(n-r) \times r} & 0_{(n-r) \times (N-r)}
\end{bmatrix}
\]

\[
\Sigma_r^{-1} = \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_r^{-1}\}
\]

\[
\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r
\]

\[
U^{-1} = U^*, \quad V^{-1} = V^*
\]

- Also called generalized inverse.
- Can be defined for rectangular and for singular matrices.
Special Cases

\[
\begin{align*}
H &= U \sum_{\mathbb{R}} V^* , \\
H^\# &= V \sum_{\mathbb{C}} U^*, \\
N \times n &= n \times n \\
N \times N &= n \times n \\
\Sigma &= \Sigma_n^{-1} \quad U^* = H^* (HH^*)^{-1} \\
\end{align*}
\]

\( r = N < n: \)

\[
H^\# = V \left[ \Sigma_n^{-1} \right] U^* = H^* (HH^*)^{-1}
\]

\( N > n = r: \)

\[
H^\# = V \left[ \Sigma_n^{-1} \quad 0_{n \times (N-n)} \right] U^* = (H^*H)^{-1} H^*
\]

\( r = N = n: \)

\[
H^\# = V \Sigma_N^{-1} U^* = H^{-1}
\]
Least Squares

\[
\begin{align*}
\min_x ||y - Ax||^2 \\
x &= A^\#y
\end{align*}
\]

- Overdetermined: no exact solution.
- Best solution in least-squares sense.
- Solution of least 2-norm.
Matrix Inversion Lemma

For $A_1$ nonsingular

\[
(A_1 + A_2 A_4^{-1} A_3)^{-1}
= A_1^{-1} - A_1^{-1} A_2 (A_4 + A_3 A_1^{-1} A_2)^{-1} A_3 A_1^{-1}
\]
Trace of a Matrix

- Sum of diagonal entries

\[ \text{tr} [ \mathbf{A} ] = \sum_{i=1}^{n} a_{ii} \]

- MATLAB

\[ \text{trace}(\mathbf{A}) \]
Trace Properties

- Trace of transpose
  \[ tr[A^T] = \sum_{i=1}^{n} a_{ii} = tr[A] \]

- Trace of sum/difference
  \[ tr[A \pm B] = \sum_{i=1}^{n} (a_{ii} \pm b_{ii}) = tr[A] \pm tr[B] \]

- Trace of product
  \[ tr[A \cdot B] = tr[B \cdot A] \]
Matrix Norm $A \in \mathbb{C}^{n \times n}$

- Satisfy norm axioms
- Froebenius Norm

$$
\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2}
$$

- Vector-induced Norm: $\|\cdot\|$ = any vector norm

$$
\|A\|_i = \sup_{\forall x \in \mathbb{C}^n} \frac{\|Ax\|}{\|x\|}
$$
Properties of Induced Norms

- **Submultiplicativity property**
  \[ \|AB\| \leq \|A\|. \|B\| \]
  \[ \|Ax\| \leq \|A\|_i. \|x\| \]

- **2-norm**
  \[ \|A\|_{i2} = \lambda_{\text{max}}^{1/2} [A^T A] \]
  \[ \|A\|_{i2} = \|A^T\|_{i2} \]