

Least-squares Solution of Homogeneous Equations

supportive text for teaching purposes

Revision: 1.2, dated: December 15, 2005

Tomáš Svoboda

Czech Technical University, Faculty of Electrical Engineering
Center for Machine Perception, Prague, Czech Republic

svoboda@cmp.felk.cvut.cz

<http://cmp.felk.cvut.cz/~svoboda>

Introduction



We want to find a $n \times 1$ vector \mathbf{h} satisfying

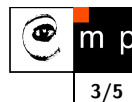
$$\mathbf{A}\mathbf{h} = \mathbf{0},$$

where \mathbf{A} is $m \times n$ matrix, and $\mathbf{0}$ is $n \times 1$ zero vector. Assume $m \geq n$, and $\text{rank}(\mathbf{A}) = n$. We are obviously not interested in the trivial solution $\mathbf{h} = \mathbf{0}$ hence, we add the constraint

$$\|\mathbf{h}\| = 1.$$

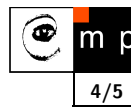
Constrained least-squares minimization: Find \mathbf{h} that minimizes $\|\mathbf{A}\mathbf{h}\|$ subject to $\|\mathbf{h}\| = 1$.

Derivation I — Lagrange multipliers



- ◆ $\mathbf{h} = \text{argmin}_{\mathbf{h}} \|\mathbf{A}\mathbf{h}\|$ subject to $\|\mathbf{h}\| = 1$. We rewrite the constraint as $1 - \mathbf{h}^T \mathbf{h} = 0$
- ◆ To find an extreme (the sought \mathbf{h}) we must solve $\frac{\partial}{\partial \mathbf{h}} (\mathbf{h}^T \mathbf{A}^T \mathbf{A} \mathbf{h} + \lambda(1 - \mathbf{h}^T \mathbf{h})) = 0$.
- ◆ We derive: $2\mathbf{A}^T \mathbf{A} \mathbf{h} - 2\lambda \mathbf{h} = 0$.
- ◆ After some manipulation we end up with: $(\mathbf{A}^T \mathbf{A} - \lambda \mathbf{E})\mathbf{h} = 0$ which is the characteristic equation. Hence, we know that \mathbf{h} is an eigenvector of $(\mathbf{A}^T \mathbf{A})$ and λ is an eigenvalue.
- ◆ The least-squares error is $e = \mathbf{h}^T \mathbf{A}^T \mathbf{A} \mathbf{h} = \mathbf{h}^T \lambda \mathbf{h}$.
- ◆ The error will be minimal for $\lambda = \min_i \lambda_i$ and the sought solution is then the eigenvector of the matrix $(\mathbf{A}^T \mathbf{A})$ corresponding to the smallest eigenvalue.

Derivation II — SVD



- ◆ Let $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where \mathbf{U} is $m \times n$ orthonormal, \mathbf{S} is $n \times n$ diagonal with descending order, and \mathbf{V}^T is $n \times n$ also orthonormal.
- ◆ From orthonormality of \mathbf{U}, \mathbf{V} follows that $\|\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{h}\| = \|\mathbf{S}\mathbf{V}^T\mathbf{h}\|$ and $\|\mathbf{V}^T\mathbf{h}\| = \|\mathbf{h}\|$.
- ◆ Substitute $\mathbf{y} = \mathbf{V}^T\mathbf{h}$. Now, we minimize $\|\mathbf{S}\mathbf{y}\|$ subject to $\|\mathbf{y}\| = 1$.
- ◆ Remember that \mathbf{S} is diagonal and the elements are sorted descendently. Than, it is clear that $\mathbf{y} = [0, 0, \dots, 1]^T$.
- ◆ From substitution we know that $\mathbf{h} = \mathbf{V}\mathbf{y}$ from which follows that sought \mathbf{h} is the last column of the matrix \mathbf{V} .

Further reading



- ◆ Richard Hartley and Andrew Zisserman, [Multiple View Geometry in computer vision](#), Cambridge University Press, 2003 (2nd edition), [Appendix A5]
- ◆ Gene H. Golub and Charles F. Van Loan, [Matrix Computation](#), John Hopkins University Press, 1996 (3rd edition).
- ◆ Eric W. Weisstein. [Lagrange Multiplier](#). From MathWorld—A Wolfram Web Resource.
<http://mathworld.wolfram.com/LagrangeMultiplier.html>
- ◆ Eric W. Weisstein. [Singular Value Decomposition](#). From MathWorld—A Wolfram Web Resource.
<http://mathworld.wolfram.com/SingularValueDecomposition.html>