FROM OBJECTIVE FUNCTIONS IN
COMPUTER VISION
TO ROBUST INLIER STRUCTURES

Peter Meer
Electrical and Computer Engineering Department
Rutgers University, NJ, USA

What I would like to accomplish in these meeting...
1. Representing nonlinear functions in a higher dimensional linear space. (Generalized) total least squares.
3. Mean shift in the Euclidean domain for segmentation and for tracking.
4. Applying nonlinear mean shift to different types of Riemannian manifold.
5. Mean shift clustering in the kernel space. Only very few cannot-link and must-link pairs are known.
6. Robust estimator for multiple inlier structures. Beside the objective function nothing else is provided.
...and will see how much I can really do.

Linear vs. nonlinear objective functions in computer vision.

There are $n$ measurements $i = 1, \ldots, n$ at the input, which can be either inliers or outliers. An unknown $n_1 \leq n$ measurements $y_i$, satisfy a $k$-dimensional vector of constraints. Given all the measurements $y_i, i = 1, \ldots, n$, find the parameters $\beta$ and number $n_1$

$$\Psi(y_i, \beta) \simeq 0 \quad i = 1, \ldots, n_1 \quad \Psi(\cdot) \in \mathbb{R}^k$$

$$\Psi(y_i, \beta) \text{ outliers} \quad i = (n_1 + 1), \ldots, n.$$ 

The case $\Psi \in \mathbb{R}^{k \times q}$ ($q > 1$) is beyond the goal of this course. The inlier points at the input, in general, satisfy a nonlinear objective function.

The $\Psi$ can be separated into a matrix of the measurements and a new parameter vector, both being derived from the input

$$\Psi(y_i, \beta) = \Phi(y_i)\theta(\beta) \quad \Phi(y) \in \mathbb{R}^{k \times m} \quad \theta(\beta) \in \mathbb{R}^m.$$ 

This is a higher dimensional linear relation if we write it as a function of carrier vectors, $x^{[c]}_i, c = 1, \ldots, \zeta$, and the new parameters $\theta$

$$x^{[c]}_i \theta \simeq 0 \quad c = 1, \ldots, \zeta \quad i = 1, \ldots, n_1.$$ 

The intercept $\alpha$ is not pulled out separately here.

A carrier contains both elements of the input measurement and pairwise products of these elements. Each of the $\zeta$ carriers are different and all these estimates must satisfy the same $\hat{\theta}$.

The $\zeta$ relations concatenated by rows is quasi-equal to $0_\zeta$ and $k = \zeta$. As will be seen later, a difference exist between $k$ and $\zeta$ when you use it for robust estimation of the objective function.

If $\zeta$ is larger that one, will take only a maximum between all $\zeta$ carriers and will do the processing in the null space with dimension $k = 1$. 

2

3

4
Fundamental matrix. $\zeta = 1$.

The inliers at the input are point correspondences from the two images $[x \ y \ x' \ y']^T \in \mathbb{R}^4$, and a $3 \times 3$, rank two, matrix $F$ has to be estimated

$$[x_1' \ y_1' \ 1] F [x_1 \ y_1 \ 1]^T \simeq 0 \quad i = 1, \ldots, n_1.$$  

The relation is not valid for points at infinity. The carrier vector is $x \in \mathbb{R}^8$

$$x = [x \ y \ x' \ y' \ xx' \ xy' \ x'y \ yy']^T$$

and the higher dimensional Euclidean space becomes

$$x_i^T \theta - \alpha \simeq 0 \quad i = 1, \ldots, n_1$$

where $n_1$ is the unknown number of inliers, and the scalar $\alpha$ (intercept) was pulled out.

Homography. $\zeta = 2$.

The inliers at the input are point correspondences between two planes in the two images $[x \ y \ x' \ y']^T \in \mathbb{R}^4$, and a $3 \times 3$ matrix $H$ has to be estimated

$$y_1' \simeq H y_1 \quad i = 1, \ldots, n_1$$

where $y = [x \ y]^T$ and $y' = [x'_i \ y_i \ u_i]^T$, not including points at infinity. The two coordinates in the second image can be written as $x' = h_i y/h_i y$ and $y' = h_i y/h_i y$ where $h_i$, is a row for the $H$ matrix.

A direct linear transformation (DLT), not pulling out $\alpha$ and $\theta = h$, is

$$A_i h = \begin{bmatrix} -y_i^T & 0_i^T & x_i'y_i^T & h_i^T \\ 0_i^T & -y_i^T & y_i'y_i^T & h_i^T \\ \end{bmatrix} \simeq 0$$

where $A_i$ is a $2 \times 9$ matrix and have to satisfy the estimated vector $h = \text{vec}(H^T)$.

The two carriers are in $\mathbb{R}^9$

$$x^{[1]} = [-x \ -y \ -1 \ 0 \ 0 \ x'x \ x'y \ x'^2]^T$$
$$x^{[2]} = [0 \ 0 \ -x \ -y \ -1 \ y'x \ y'y \ y'^2]^T$$

and in the linear space

$$x_i^{[c]} h \simeq 0 \quad c = 1, 2 \quad i = 1, \ldots, n_1.$$  

Homoscedasticity vs. heteroscedasticity

The inliers at the input $y_i$, $i = 1, \ldots, n_1$, are independent and identically distributed, having the same $l \times l$ covariance $\sigma^2 C_y$. The $\sigma^2$ is unknown and the determinant of $C_y$ is equals to one.

The covariance at the input is homoscedastic.

The covariance of the carriers are $m \times m$ matrices $\sigma^2 C^c_j$, obtained by error propagation with Jacobian matrices

$$\sigma^2 C^c_j = \sigma^2 J_{x_j y_j} C_y J_{x_j y_j}^T, \quad c = 1, \ldots, \zeta$$

where the Jacobian matrix is

$$J_{x_j y_j} = \frac{\partial x(y)}{\partial y} = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \ldots & \frac{\partial x_1}{\partial y_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_m}{\partial y_1} & \ldots & \frac{\partial x_m}{\partial y_1} \end{bmatrix}.$$  

For the carriers from nonlinear objective functions, the covariance depend on the input point from which it was computed.

In the higher dimensional linear space a different covariance exist for each of the $c = 1, \ldots, \zeta$ carriers.

The covariances of the carriers are heteroscedastic.
**Ellipse estimation.** \( \zeta = 1 \).
The inliers at the input are 2D coordinates of the points around the ellipse \( y = [x \ y]^\top \) and have to find five parameters: three distinct elements of the \( 2 \times 2 \) symmetric, positive definite matrix \( Q \) and two coordinates \( y_i \) of the ellipse center
\[
(y_i - y_i)Q(y_i - y_i) - 1 \simeq 0 \quad i = 1, \ldots, n_1.
\]
The carriers are
\[
x = [x \ y \ x^2 \ xy \ y^2]^\top
\]
and the objective function is transformed with parameters
\[
\theta = [-2y_iQ \ Q_{11} \ 2Q_{12} \ Q_{22}]^\top \quad \alpha = y_i^\top Q y_i - 1
\]
equivalent to
\[
x_i^\top \theta - \alpha \simeq 0.
\]
Besides \( \theta \theta = 1 \) the ellipses also have to satisfy the positive symmetry condition
\[
4\theta_1\theta_2 - \theta_1^2 > 0 \quad (=1).
\]
The \( 5 \times 2 \) Jacobian for the \( 5 \times 5 \) covariance of \( x \) is
\[
J_{x|y} = \begin{bmatrix}
x_{ij} & 2x_i \ y_i \\
0 & 1 & 0 & x_i & 2y_i
\end{bmatrix}
\]

Taking the second order Taylor expansion for an element of a carrier \( x_i = [\ldots \ x_{ij} \ldots]^\top \), its expected mean value can be found. With the true value having an additional \( \sigma \), we have
\[
x_{ij} \approx x_{ij} + \left[ \frac{\partial x_{ij}(y_{in})}{\partial y} \right]^\top (y_i - y_{in}) + \frac{1}{2} (y_i - y_{in})^\top \left[ \frac{\partial^2 x_{ij}(y_{in})}{\partial y^2} \right] (y_i - y_{in}).
\]
where \( \left[ \frac{\partial x_{ij}(y_{in})}{\partial y} \right] \) in the gradient.
The symmetric \( l \times l \) Hessian matrix is \( H_{ij}(y_i) = \sigma^2 \frac{\partial^2 x_{ij}}{\partial y^2} \)
and using \( y \) instead of \( y_{in} \) gives the expectation
\[
E[x_{ij} - x_{ij}] = \sigma^2 \frac{\partial^2 x_{ij}}{\partial y^2}.
\]
Most carriers do not have second order terms derived from the input, and the expected means are zero.

The **ellipses** are different. The \( 2 \times 2 \) Hessians are
\[
H_1 = H_2 = 0 \quad H_3 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad H_5 = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.
\]
For \( \sigma^2 C_y = \sigma^2 I_2 \) the carriers have
\[
E[x - x_i] = \begin{bmatrix} 0 & 0 & \sigma^2 & 0 & \sigma^2 \end{bmatrix}^\top.
\]

**Fundamental matrix.** \( \zeta = 1 \).
The \( 8 \times 4 \) Jacobian matrix for the \( 8 \times 8 \) covariance of \( x \), with the \( \alpha \) pulled out, is
\[
J_{x|y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\]

**Homography.** \( \zeta = 2 \).
The two \( 9 \times 4 \) Jacobian matrices for the \( 9 \times 9 \) covariances, without \( \alpha \) pulled out, are
\[
J_{x|y} = \begin{bmatrix} -I_{2 \times 2} & x_i^\top I_{2 \times 2} & 0_2 \\ 0_2 & 0_{1 \times 4} \end{bmatrix}, \quad J_{x|y} = \begin{bmatrix} 0_{1 \times 3} & 0_1 & 0_2 & 0_2 \\ 0_2 & y_i^\top & 0_2 \end{bmatrix}
\]
All the covariances depend on the input point from where they were derived.

\[
n_1 = n \quad \text{nonrobust estimation}
\]
Assume the surface \( g(y_n) = 0 \) and is not very nonlinear. The noise is \( \delta y_i \sim G(0, \sigma^2 I) \). The \( g(y) \) is not equal to zero and the first order Taylor expansion is
\[
g(y) \approx g(\hat{y}) + [\nabla_y g(\hat{y})]^\top (y - \hat{y}) = [\nabla_y g(\hat{y})]^\top (y - \hat{y})
\]
where \( g(\hat{y}) = 0 \) and \( \nabla_y g(\hat{y}) \) is the gradient of \( g(\hat{y}) \).
Taking the norms in \( L_2 \)
\[
|g(y)|_2 \approx ||\nabla_y g(\hat{y})||_2 ||y - \hat{y}||_2
\]
and using \( y \) instead of \( \hat{y} \)
\[
||y - \hat{y}||_2 \approx \frac{|g(y)|_2}{||\nabla_y g(\hat{y})||_2}
\]
gives the **algebraic distance** from \( y \). In general is not geometrically meaningful, but works as initial solution.
If the objective function is linear, e.g.,
\[
g(y) = y_i \quad \theta - \alpha \approx 0 \quad \theta' \theta = 1 \quad i = 1, \ldots, n
\]
the gradient is \( \nabla_y g(\hat{y}) = \theta \) and the norm for the gradient is one
\[
||y - \hat{y}|| = |g(y)|_2.
\]
Will take \( y \) and \( \theta \) vectors with dimension \( m \).
The linear objective function with the simplest linear errors-in-variables (EIV) regression model is

\[ g(y_i) = y_i^\top \theta - \alpha = 0 \]

The smallest, right singular vector of

\[ \tilde{\alpha}, \hat{\theta} = \arg\min_{\alpha, \theta} \| y_i - \tilde{\alpha} \|^2 \] 

subject to \( y_i^\top \tilde{\alpha} = \alpha \) and \( \hat{\theta}^\top \hat{\theta} = 1 \).

The data is centered using the \( n \times n \) orthogonal projection matrix

\[ G = I_n - \frac{1}{n} 1_n 1_n^\top \]

to obtain \( \hat{Y} \) and \( \tilde{\theta} \). Example: \( \n \in \in \) \( \in \)

The centroid of the data is on the TLS solution.

The full rank \( m \times m \) covariance is given by the user \( \sigma^2 C_y \). Estimate the minimization

\[ \hat{\alpha}, \hat{\theta} = \arg\min_{\alpha, \theta} \| y_i - \hat{\alpha} \|^2 \] 

subject to \( y_i^\top \hat{\alpha} = \alpha \) and \( \hat{\theta}^\top \hat{\theta} = 1 \).

The gradient in \( \hat{\theta} \), equal to zero

\[ \frac{\partial J_{\text{TLS}}}{\partial \hat{\theta}_i} = 0 \quad \Rightarrow \quad \hat{\theta}_i = \hat{y}_i - \eta_i C_y \hat{\theta} \]

and multiplying both sides with \( \hat{\theta}^\top \), we obtain

\[ \eta_i = \frac{\hat{\theta}^\top \hat{y}_i}{\hat{\theta}^\top C_y \hat{\theta}} \]

Thus

\[ \hat{y}_i - \hat{\tilde{y}}_i = \frac{\hat{\theta}^\top \hat{y}_i C_y \hat{\theta}}{\hat{\theta}^\top C_y \hat{\theta}} \]

and \( J_{\text{TLS}} \) can be written

\[ J_{\text{TLS}} = \sum_{i=1}^{n} (\hat{\theta}^\top C_y \hat{\theta})(\hat{y}_i - \hat{\tilde{y}}_i)^\top C_y^{-1} (\hat{y}_i - \hat{\tilde{y}}_i) + \sum_{i=1}^{n} \eta_i (\hat{y}_i - \hat{\tilde{y}}_i)^\top \hat{\theta} \]

where each term in the second part is equal to zero.
The gradient in \( \hat{\theta} \) is also equal to zero \( \frac{\partial J_{GTLS}}{\partial \hat{\theta}} = 0 \).

Gives

\[
\frac{2Y^\top \hat{Y} \hat{\theta} (\hat{\theta}^\top C_y \hat{\theta}) - 2(\hat{\theta}^\top Y^\top Y \hat{\theta}) C_y \hat{\theta}}{(\hat{\theta}^\top C_y \hat{\theta})^2} = 0
\]

or

\[
Y^\top \hat{Y} \hat{\theta} = \lambda C_y \hat{\theta} \quad \lambda = \frac{\hat{\theta}^\top Y^\top Y \hat{\theta}}{\hat{\theta}^\top C_y \hat{\theta}}
\]

which is a generalized eigenvalue problem, minimized by \( \lambda \) taking the smallest eigenvalue.

The condition \( \hat{\theta}^\top \hat{\theta} = 1 \) is imposed at the end.

The TLS and the generalized TLS were solved in different ways? Not really. If \( C_y = I_m \) then

\[
\begin{align*}
\text{GTLS} & \quad Y^\top \hat{Y} \hat{\theta} = \lambda \hat{\theta} \\
\text{TLS} & \quad \hat{Y} \hat{\theta} = 0
\end{align*}
\]

and the eigenvector corresponding to the smallest eigenvalue \( \lambda_{\min} \) of \( Y^\top Y \), is the same with the vector spanning the null space of \( \hat{Y} \). Both are equal to \( v_m \).