Lecture L5

Today:

1. Gaussian Markov random fields and sparse matrices
2. Soft Constraints
3. Intrinsic Gaussian Markov random fields
4. Matlab tricks
5. Example: Pressure field over Europe

Exercise

We have observed almost all pixels of an image $\boldsymbol{x}$, with known expectation $\mu = 5$:

$$\boldsymbol{x} = \begin{bmatrix} 4 & 3 & 5 & 1 \\ 2 & 8 & ? & 6 \\ 3 & 5 & 7 & 6 \\ 3 & 6 & 5 & 4 \end{bmatrix}$$

We know (somehow) that the row of the precision matrix corresponding to the missing pixel is

$$\boldsymbol{Q}_{10,\cdot} = [0, 0, 0, 0, 0, -1, 0, 0, -3, 10, -3, 0, 0, -1, 0, 0]$$

Calculate the expected value of the missing pixel, $x_{2,3}$, given all the other pixels.

Exercise solution

```
mu: 5
x: [4  3  5  1;      x-mu: [-1 -2  0 -4;
    2  8  -  6;             -3  3  -  1;
    3  5  7  6;             -2  0  2  1;
    3  6  5  4]             -2  1  0 -1]
icolstack(Q(10,:)',[4,4]):
   [0  0 -3  0;     c: [0     0 3/10    0;
    0 -1 10 -1;         0  1/10    0 1/10;
    0  0 -3  0;         0     0 3/10    0;
    0  0  0  0]         0     0    0    0]
E(x(2,3)): mu + sum(sum(c.*(x-mu))) = 5 + (0*3+3+1+2*3)/10
Conditional variance = 1/10
```

(The model does not really seem like a good model; is it realistic to get 16 integers in a Normal distribution, especially with conditional standard deviation $\tau \approx 0.3$?)

Sparse matrices and GMRF

Last lecture we introduced pixel dependencies by modeling the entire image using a large Gaussian distribution, $\boldsymbol{X} \in \mathsf{N}(\mu, \Sigma)$. The problem here is that Σ is a dense $mn \times mn$-matrix!
But:

- If $\boldsymbol{X}$ is a Markov field the precision matrix $\boldsymbol{Q} = \Sigma^{-1}$ will be sparse.
- If the field is homgenous $\boldsymbol{Q}$ is highly structured and can be defined using a precision function $q(\boldsymbol{u})$.
- Under mild conditions the Cholesky factorisation of $\boldsymbol{Q} = (\boldsymbol{R}^\mathsf{T}\boldsymbol{R})$, will be sparse if $\boldsymbol{Q}$ is sparse.
- However $\boldsymbol{R}^{-1}$ is a dense matrix.

GMRFs become computationally feasible if we replace operations that use Σ with operations using $\boldsymbol{Q}$ or $\boldsymbol{R}$.

Calculation of conditional expectation

The conditional distribution for some pixels given the rest,

$$\boldsymbol{X}|\boldsymbol{Y} \in \mathrm{N}(\mu_{\boldsymbol{X}} - \boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{X}}^{-1}\boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{Y} - \mu_{\boldsymbol{Y}}), \boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{X}}^{-1})$$

We want to calculate $\mathrm{E}(\boldsymbol{X}|\boldsymbol{Y})$ without having to calculate the dense, potentially large, matrix $\boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{X}}^{-1}$.

If $\boldsymbol{Q} = \boldsymbol{R}^\top\boldsymbol{R}$ then the Cholesky factorisation of $\boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{X}}$ is $\boldsymbol{R}_{\boldsymbol{X},\boldsymbol{X}}^\top\boldsymbol{R}_{\boldsymbol{X},\boldsymbol{X}}$ and the conditional expectation can be calculated as,

$$\mathrm{E}(\boldsymbol{X}|\boldsymbol{Y}) = \mu_{\boldsymbol{X}} - \boldsymbol{R}_{\boldsymbol{X},\boldsymbol{X}}^{-1}\left(\boldsymbol{R}_{\boldsymbol{X},\boldsymbol{X}}^{-\top}\boldsymbol{Q}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{Y} - \mu_{\boldsymbol{Y}})\right),$$

using back-substitution.

Matrix inversion lemma

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1} + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{V}\boldsymbol{A}^{-1}$$

If $\boldsymbol{A}^{-1}$ is known and $\boldsymbol{C}$ is much smaller than $\boldsymbol{A}$ the lemma gives an efficient way of obtaning $(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1}$ if $\boldsymbol{A}^{-1}$ is known. It can also be used to simplify expressions that arise when dealing with Gaussian Markov Fields.

Soft constraints

Assume that we have a GMRF, $\tilde{\boldsymbol{X}} \in \mathrm{N}(\mu, \boldsymbol{Q}^{-1})$, and that we have observed some linear combinations of the field with errors $\tilde{\boldsymbol{Y}} = \boldsymbol{A}\tilde{\boldsymbol{X}} + \tilde{\boldsymbol{E}}$, where $\tilde{\boldsymbol{E}} \in \mathrm{N}(\boldsymbol{0}, \Sigma_{\boldsymbol{E}})$.

- The conditional precision is, $\boldsymbol{Q}_{\boldsymbol{X}|\boldsymbol{Y}} = \boldsymbol{Q} + \boldsymbol{A}^\top{\Sigma_{\boldsymbol{E}}}^{-1}\boldsymbol{A}$
- The conditional distribution is,

$$\boldsymbol{X}|\boldsymbol{Y} \in \mathrm{N}\left(\boldsymbol{Q}_{\boldsymbol{X}|\boldsymbol{Y}}^{-1}(\boldsymbol{Q}\mu + \boldsymbol{A}^\top{\Sigma_{\boldsymbol{E}}}^{-1}\boldsymbol{Y}), \boldsymbol{Q}_{\boldsymbol{X}|\boldsymbol{Y}}^{-1}\right),$$

  which can be handled efficiently if $\boldsymbol{Q}_{\boldsymbol{X}|\boldsymbol{Y}}$ is sparse.
- Applying the matrix inversion lemma to $\boldsymbol{Q}_{\boldsymbol{X}|\boldsymbol{Y}}^{-1}$ the conditional expectation can be rewritten as

$$\mu_{\boldsymbol{X}|\boldsymbol{Y}} = \mu - \boldsymbol{Q}^{-1}\boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^\top + \Sigma_E)^{-1}(\boldsymbol{A}\mu - \boldsymbol{Y}),$$

  which is computationaly feasible if $K \ll N$.

Soft constraints (cont)

Efficient calculation of:

$$\mu_{\boldsymbol{X}|\boldsymbol{Y}} = \mu - \boldsymbol{Q}^{-1}\boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^\top + \Sigma_E)^{-1}(\boldsymbol{A}\mu - \boldsymbol{Y}).$$

1. Compute $\boldsymbol{R}$ such that $\boldsymbol{Q} = \boldsymbol{R}^\top\boldsymbol{R}$.
2. Calculate $\boldsymbol{B} = \boldsymbol{Q}^{-1}\boldsymbol{A}^\top = \boldsymbol{R}^{-1}(\boldsymbol{R}^\top)^{-1}\boldsymbol{A}^\top$ through back-substitution.
3. Compute $\boldsymbol{R}_2$ such that $\boldsymbol{A}\boldsymbol{B} + \Sigma_E = \boldsymbol{R}_2^\top\boldsymbol{R}_2$.
4. Compute $\boldsymbol{C} = \boldsymbol{B}(\boldsymbol{A}\boldsymbol{B} + \Sigma_E)^{-1} = \boldsymbol{B}\boldsymbol{R}_2^{-1}(\boldsymbol{R}_2^\top)^{-1}$ through back-substitution.
5. Compute $\mu_{\boldsymbol{X}|\boldsymbol{Y}} = \mu - \boldsymbol{C}(\boldsymbol{A}\mu - \boldsymbol{Y})$.

Note that all the stored dense matrices are much smaller than $\boldsymbol{Q}$.

Soft constraints (cont)

Having found the conditional expectation we are also interested in the pointwise conditional variances $\mathsf{V}(X_{\boldsymbol{u}}|\boldsymbol{Y})$.

- The pointwise variances are given by the diagonal elements of $\boldsymbol{Q}_{X|Y}^{-1}$, which can be calculated without inverting the entire matrix.
- Simulate from $\boldsymbol{X}|\boldsymbol{Y}$ and calculate empirical variances.
  1. Simulate $\boldsymbol{z}$ from $\mathsf{N}(\boldsymbol{Y}, \Sigma_E)$.
  2. Simulate $\boldsymbol{m}$ from $\mathsf{N}(\mu, \boldsymbol{Q}^{-1})$.
  3. Now $\boldsymbol{m} - \boldsymbol{C}(\boldsymbol{Am} - \boldsymbol{z})$ is a sample from $\tilde{\boldsymbol{X}}|\boldsymbol{Y}$.

The Markov condition for image MRF:s

- Recall that a random field which fulfils the relation:

$$p(x_{\boldsymbol{u}}|x_{\boldsymbol{v}}, \boldsymbol{v} \neq \boldsymbol{u}) = p(x_{\boldsymbol{u}}|x_{\boldsymbol{v}}, \boldsymbol{v} \in \mathcal{N}_{\boldsymbol{u}})$$

  for some system of neighbourhoods $\mathcal{N} = \{\mathcal{N}_{\boldsymbol{u}}\}$ is called a *Markov random field*.
- Can we specify global models via the local, conditional distributions?

Using increments

If we want a smooth field it seems reasonable to assume that the increments between neighbouring pixels are Normal and "independent".

- This gives, $x_{i,j} - x_{i,j-1} \in \mathsf{N}(0, \varkappa^{-1})$ and $x_{i,j} - x_{i-1,j} \in \mathsf{N}(0, \varkappa^{-1})$.
- Gathering the increments in a matrix $\boldsymbol{W}$ where each row corresponds to one increment we obtain $\boldsymbol{WX} \in \mathsf{N}(\boldsymbol{0}, \varkappa^{-1}\boldsymbol{I})$.
- The density of $\boldsymbol{WX}$ is

$$p(\boldsymbol{WX}) \propto \exp\left(\frac{-1}{2}(\boldsymbol{WX})^{\mathsf{T}}(\boldsymbol{I}/\varkappa)^{-1}(\boldsymbol{WX})\right)$$

- Simplifying $p(\boldsymbol{WX})$ the density for the entire field is $\mathsf{N}(\boldsymbol{0}, (\varkappa\boldsymbol{Q})^{-1})$ with $\boldsymbol{Q} = \boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}$

Using increments, an example

For a 3×3 image the increment matrix is

$$\boldsymbol{W} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix},$$

and some selected rows of the $\boldsymbol{Q}$-matrix are

$$\boldsymbol{Q}_1 = \begin{pmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\boldsymbol{Q}_2 = \begin{pmatrix} -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\boldsymbol{Q}_5 = \begin{pmatrix} 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \end{pmatrix}$$

Intrinsic Gaussian Markov random fields (IGMRF)

Studying the $\boldsymbol{Q}$-matrix we see that $\boldsymbol{Q}\mathbf{1} = \mathbf{0}$. This implies that:

- The field is invariant to the addition of a constant since $p(\boldsymbol{X} + c\mathbf{1}) = p(\boldsymbol{X})$
- The density $p(\boldsymbol{X})$ is *improper* since it can not be integrated.
- The determinant $|\boldsymbol{Q}|$ has to be replaced with the *generalised determinant* $|\boldsymbol{Q}|_*$
- Since $\mathbf{1}$ is the only eigenvector with eigenvalue 0 $\text{rank}(\boldsymbol{Q}) = n - 1$
- $|\varkappa\boldsymbol{Q}|_* = \varkappa^{n-1}|\boldsymbol{Q}|_*$

IGMRF (cont)

$\varkappa\boldsymbol{Q} + \gamma\mathbf{1}\mathbf{1}^\mathsf{T}$ is a matrix of full rank which defines a proper density. Using this precision matrix the density becomes,

$$\begin{aligned} p(\boldsymbol{X}) &\propto \exp\left(-\boldsymbol{X}^\mathsf{T}\left(\varkappa\boldsymbol{Q} + \gamma\mathbf{1}\mathbf{1}^\mathsf{T}\right)\boldsymbol{X}/2\right) \\ &= \exp\left(-\boldsymbol{X}^\mathsf{T}(\varkappa\boldsymbol{Q})\boldsymbol{X}/2\right)\exp\left(-\boldsymbol{X}^\mathsf{T}\gamma\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{X}/2\right) \\ &= \exp\left(-\boldsymbol{X}^\mathsf{T}(\varkappa\boldsymbol{Q})\boldsymbol{X}/2\right)\exp\left(-\gamma n^2\overline{\boldsymbol{X}}^2/2\right). \end{aligned}$$

- The first part coresponds to the IGMRF.
- The second part can be seen as a prior pulling the expectation of the field towards 0, with $\text{V}(\overline{\boldsymbol{X}}) = 1/(\gamma n^2)$.
- The IGMRF is obtained by letting $\gamma \to 0$

2:nd order IGMRF

The 1:st order IGMRF is obtained by assuming that the first order "derivatives" are Normal. If we instead let the second order "derivatives" be Normal we obtain a 2:nd order IGMRF.

- The 2:nd order "derivaties" are the discrete approximations to laplace.
- A 2:nd order field is "smoother" than a 1:st order field.
- The neigbourhood, $\mathcal{N}$, of a 2:nd field is larger than for a 1:st order field.

2:nd order IGMRF (cont)

The 2:nd order IGMRF is defined using the increment

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

with edge increments $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^\mathsf{T}$ and corner increment,

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

- The field is invariant to the addition of a plane of arbitrary slope.
- $\text{rank}(\boldsymbol{Q}) = n - 3$
- $|\varkappa\boldsymbol{Q}|_* = \varkappa^{n-3}|\boldsymbol{Q}|_*$

The order of an IGMRF

There are several ways of defining the order of a IGMRF,

- Order based on the rank deficiency.
- order based on the properties of the local increments.
- If the field is invariant to the addition of a all polynomials with degree less than or equal to $k-1$ the (polynomial) order of the field is k.

Matlab

Simulation of Gaussian images.

```
>> Q = Precision matrix for image of size m-by-n;
>> R = chol(Q);      % Q = R'*R
>> Z = randn(m*n,1); % Z i.i.d. N(0,1)
>> X = mu + R\Z;     % Note: Not "inv(R)"!
>> x = icolstack(X,[m,n]);
```

- Given: $\boldsymbol{Z} \in \mathsf{N}(\boldsymbol{0}, \boldsymbol{I})$, and $\boldsymbol{R}$ such that $\boldsymbol{Q} = \boldsymbol{R}^\top \boldsymbol{R}$.

$$
\begin{aligned}
\mathsf{E}(\boldsymbol{X}) &= \mathsf{E}(\mu + \boldsymbol{R}^{-1}\boldsymbol{Z}) = \mu + \boldsymbol{R}^{-1}\mathsf{E}(\boldsymbol{Z}) = \mu \\
\mathsf{Cov}(\boldsymbol{X},\boldsymbol{X}) &= \mathsf{E}((\boldsymbol{X}-\mu)(\boldsymbol{X}-\mu)^\top) = \mathsf{E}((\boldsymbol{R}^{-1}\boldsymbol{Z})(\boldsymbol{R}^{-1}\boldsymbol{Z})^\top) \\
&= \boldsymbol{R}^{-1}\mathsf{E}(\boldsymbol{Z}\boldsymbol{Z}^\top)\boldsymbol{R}^{-\top} = \boldsymbol{R}^{-1}\boldsymbol{I}\boldsymbol{R}^{-\top} = (\boldsymbol{R}^\top\boldsymbol{R})^{-1} = \boldsymbol{Q}^{-1}
\end{aligned}
$$

Matlab tricks

Logical indexing:

```
>> x = [3 4 1;
        5 9 6;
        7 2 8];
>> k = logical([0 1 1;
                1 0 1;
                1 1 0]);
>> x(k(:))'
  [5 7 4 2 1 6]
>> x(~k(:))'
  [3 9 8]
```

```
>> k = logical([1 0 1]');
>> x(~k,:)
  [5 9 6]
```


More Matlab

```
X = colstack(x);
Q = gmrfprec(size(x),[0 -1 0;-1 5 -1;0 -1 0]);
spy(Q)
K = logical(known(:));
Q_rearranged = [Q(K,K),Q(K,~K);Q(~K,K),Q(~K,~K)];
spy(Q_rearranged)
Z = zeros(size(X));
Z(K) = x(K);
Z(~K) = mu(~K) - Q(~K,~K)\(Q(~K,K)*(X(K)-mu(K)));
z = icolstack(Z,size(x));
```

$\boldsymbol{z}$ is now the conditional expectation of the "unknown" pixels, given the "known" pixels.

Further reading

- Rue & Held, *Gaussian Markov Random Fields, Theory and Applications*
- GMRFLib a free C library for GMRF calculations, `http://www.math.ntnu.no/~hrue/GMRFLib/`

An example

- We have pressure measurements in millibar for several European weather stations from $2006 - 08 - 31$, $18 : 00$, and want to construct a pressure field for Europe.
- Use a soft constrain model on a 2:nd order IGMRF:

$$\tilde{\boldsymbol{X}} \in \mathsf{N}\left(0, (\varkappa \boldsymbol{Q})^{-1}\right)$$
$$\tilde{\boldsymbol{Y}} | \boldsymbol{X} \in \mathsf{N}\left(\boldsymbol{A}\boldsymbol{X}, \sigma^2 \boldsymbol{I}\right).$$

- The $\boldsymbol{A}$-matrix is constructed by checking in which "pixel" each weather station lies.
- The field can then be reconstructed using the formula for conditional expectation for soft constraints.
- The pointwise variance is found through simulation.
- Note that standard pressure is 1013.25 mb